VOLUME 26 • NUMBER 4

December 1955

the RESEARCH quarterly

STEVENS RICE
UNIVERSITY MICROFILMS
313 N FIRST ST
ANN ARBOR MICH
HPR DEC P ADV

Published by THE AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION
1201 SIXTEENTH ST., N. W., WASHINGTON 6, D. C.

DON'T OVERLOOK THESE RECENT PUBLICATIONS!

Dance Research and Theses: Completed, In Process, with Research Needed

Available from Chairman of Dance Section Research Committee: Virginia Moomaw, Dept. of Physical Education, Woman's College, University of North Carolina, Greensboro, N. C.

Mimeo, 57 pp. $1.00

A Directory of Colleges and Universities Offering a Major, Minor or Concentration Dance Curriculum

Compiled by National Dance Section of AAHPER.

Mimeo, 24 pp. 50c

Physical Education for College Men and Women

Complete report of the Washington Conference, Oct. 1954, sponsored by AAHPER, CPEA, and NAPECW.

48 pp. $1.00

Health Education Conference Report

Report of the National Conference on the Undergraduate Health Education Minor Program and Desirable Health Education Emphases for the Physical Education Major Program, Washington, Jan. 1955.

36 pp. $1.00

Professional Contribution No. 3

Papers and reports of two annual meetings, American Academy of Physical Education, Chicago, 1953; New York City, 1954.

Includes four papers on the effects of competition in interschool athletics.

146 pp. $1.50

Order from

AAHPER Publications-Sales
1201 - 16th St., N.W.,
Washington 6, D. C.

the RESEARCH Quarterly

of the American Association for Health, Physical Education, and Recreation, Department of the National Education Association, 1201 Sixteenth St., N.W., Washington 6, D. C.

Vol. 26, No. 4 DECEMBER 1955

CONTENTS

Attitude of University of Washington Women Students Toward Physical Education Activity ..... MARION R. BROER, KATHARINE S. FOX, and EUNICE WAY 379

Objective Determination of Resistance Load for Ten Repetitions Maximum for Quadriceps Development
LT. DAVID H. CLARKE, MSC and LT. EDWARD L. HERMAN, MSC 385

Influence of Wheat Germ Oil as a Dietary Supplement in a Program of Conditioning Exercises with Middle-Aged Subjects
THOMAS KIRK CURETON, assisted by RICHARD H. POHNDORF 391

Instruments To Measure Strength PAUL A. HUNSICKER and RICHARD J. DONNELLY 408

Aggression in Boxers and Wrestlers As Measured by Projective Techniques
BURRIS F. HUSMAN 421

Relationship Between Strength Loss in Pack Carrying and Certain Motor-Physical Fitness Criteria
DONALD K. MATHEWS, CLAYTON T. SHAY, and H. HARRISON CLARKE 426

Comparative Analysis of Methods of Scoring Tests of Motor Learning
L. W. McCRAW 440

Cardiovascular Measurements on College-Age Smokers and Non-Smokers
NICHOLAS PETER MOUTIS 454

Health and Physical Education Program for Boys in Ohio High Schools
MAJOR LAURENCE E. OWENS, USAF 461

Comparisons of Reaction-Time Measures to a Visual Stimulus and Arm Movement
A. T. SLATER-HAMMEL 470

NOTES AND COMMENTS—**Russian Studies of Vision in Relation to Physical Activity and Sports, Part I,** Ashton Graybiel, M.D., Ernst Jokl, M.D., and Claude Trapp, M.D., **480; Swimming Classification Test,** M. Gold and I. F. Waglow, **485; Probability as Applied to the Design of Experiments in Health, Physical Education, and Recreation,** Henry J. Montoye, **486.**

Research Abstracts 488

Life and Honorary Members 492

Representative Assembly of AAHPER, 1955-56 495

Index for Volume 26 497

Copyright 1955, by the American Association for Health, Physical Education and Recreation National Education Association

Carl A. Troester, Jr., Executive Secretary

Assistant Secretaries

Elizabeth Avery, Consultant in Health Education

Rachel E. Bryant, Consultant in Physical Education and Girls and Women's Sports

Jackson M. Anderson, Consultant in Recreation and Outdoor Education

Eugene H. Sloane, Director of Publications and Public Relations

●

Ella H. Wright, Editor

Jeanette K. Cokouros, Assistant to the Editor

Jean W. Sloan, Circulation Manager

George F. Anderson, Associate Executive Secretary in charge of Advertising

Ames A. Castle, Sports and Industrial Relations Executive

●

Board of Associate Editors

Karl Bookwalter
Marion R. Broer
H. Harrison Clarke
Thomas K. Cureton, Jr.
Anna Espenschade
Franklin M. Henry
Pauline Hodgson
James H. Humphrey
Paul Hunsicker
Leslie W. Irwin
Warren R. Johnson
H. F. Kilander
Joy W. Kistler
Hyman Krakower
Lynn W. McCraw
Kenneth D. Miller
Henry J. Montoye
Laurence E. Morehouse
Marjorie Phillips
Laurence Rarick
M. Gladys Scott
John H. Shaw
Frank D. Sills
Wesley M. Staton
D. B. Van Dalen
C. Etta Walters
Allen E. Weatherford

Published in March, May, October, and December by the American Association for Health, Physical Education, and Recreation, 1201 Sixteenth Street, N.W., Washington 6, D. C. Subscription available to libraries, institutions, and departments at $5.00 per year. Of the Professional Membership ($10.00 per year), $3.00 covers subscription to the *Quarterly*. Of the Student Professional Membership ($5.00 per year), $1.50 is for subscription to the *Quarterly*. *Subscriptions to individuals through membership only.* Single copies $1.25.

Send membership dues or subscriptions to Circulation Dept., 1201 Sixteenth Street, N.W., Washington 6, D. C. Editorial Office, 1201 Sixteenth Street, N.W., Washington 6, D. C. Entered as second-class matter at the Post Office at Washington, D. C., under the act of March 3, 1879. Additional entry at Baltimore, Md.

Manuscripts are to be sent to the Editorial Office and are expected to follow the form prescribed in the Guide to Authors that appears each year in the October issue of the Research Quarterly.

Attitude of University of Washington Women Students Toward Physical Education Activity

MARION R. BROER, KATHARINE S. FOX AND EUNICE WAY
University of Washington
Seattle, Washington

Abstract

This article presents the results of a survey of the attitude of 1,149 college Freshman women toward physical education as an activity. Study of the total scores shows that they indicated a very favorable attitude. The students in swimming and tennis seem to have a more favorable attitude than average and those in archery a less favorable attitude. The high percentages of these students who indicated that physical education activity classes contribute to social development, mental and physical health agree with results found at the University of Michigan.

THE PURPOSE OF this study was to determine University of Washington Freshman and Sophomore women students' attitude toward physical education activity as measured by the Wear Physical Education Attitude Inventory (4).

Procedure

Two weeks prior to the end of spring quarter, 1953, the Wear Attitude Inventory was administered to 1,149 Freshman and Sophomore women students who were at that time enrolled in physical education activity classes. The inventory was administered in the activity classes and no attempt was made to secure data on students absent from class on the day that it was given.

The inventories were scored according to directions and the median, mean, standard deviation, and sigma of the mean of the scores for the entire group were calculated. The percentages of students falling within the decidedly favorable, favorable, undecided, unfavorable, and decidedly unfavorable scores were calculated.

The papers were divided according to activity classes in which the students were enrolled at the time and according to instructors teaching the classes. The mean attitude scores, the standard deviations, and sigmas of the means of students enrolled in each activity and being taught by each instructor were calculated. Differences were tested for significance by the "t" test.

From the various activities and instructors, groups of 100 inventories were selected at random. The numbers of students who responded as strongly favorable, favorable, undecided, unfavorable and decidedly unfavorable to each statement were tabulated and the percentages calculated for the first 100, and then for the second 100 inventories. These results were compared. Similar results were apparent. A third 100 inventories were then tabulated

and the results again compared with the first and second hundreds. The pattern of results was consistent again and, therefore, these 300 were taken as a sample for further study. To test the sample further, the total inventory scores of the sample of 300 were compared to those of the entire sample of 1,149. It was found that the mean of the sample was slightly but not significantly higher than that of the total group. These figures were found to be:

	Total group	Sample
Number	1,149	300
Mean	150.5	153
Median	155	157
Sigma	24.6	23.5
Range	60 - 198	65 - 193

Each item of the inventory was studied by comparing the percentage of students who checked a favorable response with the percentage of students who checked an unfavorable response. Two comparisons were made: 1. the percentage who were strongly favorable with that decidedly unfavorable; 2. the percentage whose responses indicated either a strongly favorable or favorable attitude with the percentage indicating either an unfavorable or decidedly unfavorable attitude. The sigmas of the differences in the percentages were calculated and the differences were tested for significance.

Discussion of Results

Scores on the inventory indicate attitude toward physical education as follows:

200—totally[1] strongly favorable attitude
160—totally favorable attitude
120—completely neutral (undecided) attitude
80—totally unfavorable attitude
40—totally decidedly unfavorable attitude

The mean score for 1,149 students is found to be 150.5 which approaches the totally favorable score. The median is 155. The percentages of scores that fall at various levels are:

180 – 200 – 9%
160 – 179 – 27%
140 – 159 – 35%
120 – 139 – 17%
100 – 119 – 9%
Below 100 – 3%

These figures indicate that 36 per cent of these students responded with a completely positive attitude toward physical education activity; 52 per cent were undecided to completely positive; and only 12 per cent were undecided to negative. Only 3 per cent were below 100, which is halfway between the undecided and the unfavorable score; and of these, only 0.3 per cent responded with totally unfavorable attitudes. None was completely decidedly

[1]On all 40 items of the inventory.

unfavorable. If scores from 100 to 140 are considered neutral, it is found that 71 per cent are favorable as opposed to 3 per cent unfavorable to physical education activity.

While it is obvious that the attitude expressed by a student will be influenced by all of her past experiences, it was reasoned that the most recent physical education experience would be likely to have a strong effect. Reeves (3, *p. 300*), in a study at the University of Southern California, concluded that students were more influenced by the class in which they were currently enrolled than by any other factor. Therefore, it was felt that it would be of interest to determine whether differences in mean attitude scores of those enrolled in various activities and of those taught by various instructors were apparent.

When the mean attitude scores of the various activity groups are studied, it is found that those enrolled in swimming at the time that the inventory was administered indicated the most favorable attitude toward physical education activity (see Table 1). Tennis is second. The lowest means are found in the badminton and archery groups. Since at the University of Washington badminton is a much more popular activity in winter than in spring, it may be that students who were unable, for various reasons, to enroll in an outdoor activity of their choice were in these classes and this may in some degree account for the showing made by this group. Also, the group involved is small. However, spring is the most popular quarter for archery and the group is large enough that the results probably indicate that some of the students in these classes do not have the most favorable attitude toward physical education activity. While many students are very enthusiastic about archery, it has long been recognized that this is an activity in which students unfavorable to physical education may enroll because they do not have to

TABLE 1

Mean Attitude Scores of Students Enrolled in Various Activities

Activity	No.	Mean	σ	σ_M
Archery	97	142.95	24.5	2.50
Badminton	28	140.68	25.4	4.79
Bowling	140	146.75	25.0	2.12
Canoeing	123	149.99	23.0	2.07
Fencing	44	148.34	31.0	4.70
Folk dance	89	149.57	22.0	2.34
Golf	91	146.46	22.5	2.37
Modern dance	37	148.95	24.0	3.93
Riding	16*	143.0		
Softball	12*	153.0		
Swimming	139	158.83	19.5	1.65
Tap dance	28	149.61	23.5	4.43
Tennis	305	152.82	21.5	1.23

*Further statistics were not calculated for groups under 20

change clothes and they feel it will require less physical energy. It must be noted that weather may be a factor. At the University of Washington, there are no facilities for indoor shooting and therefore, rainy-day lessons must depend upon lecture, discussions, and visual aids material. It is interesting that in a study of low motor ability students at the University of Washington (2, *p. 21*) it was found that the improvement of the attitude of *low* motor ability students during a quarter of instruction in archery was surpassed only by their attitude improvement during swimming and modern dance.[2] This is no doubt due to the low motor ability students being more able to compete in archery than in some of the other activities.

A study of the significance of the differences between these means of various activity groups shows that the mean of the group enrolled in various swimming classes is significantly higher than that of groups in other activities.[3] Between swimming and archery, badminton, bowling, canoeing, golf, folk dancing and tennis, the differences are significant at the 1-per-cent level of confidence; between swimming and modern dance, at the 2-per-cent level; and between swimming and fencing, at the 5-per-cent level. These results are based on the total swimming group. They embrace eight different levels of swimming instruction; beginning, elementary, intermediate, advanced, life saving, water safety instructors, rhythmic swimming, and limited swimming are reported as a whole.

The mean attitude score of those enrolled in tennis is significantly higher than that of those taking archery (1-per-cent level of confidence); and badminton, bowling and golf (2-per-cent level). There were classes in both intermediate and beginning tennis. While the mean attitude score of the intermediate group (No. = 41; Mean = 155.7) is higher than that of the beginning group (No. = 264; Mean = 151.8), it is not significantly so. The canoeing group's mean score is significantly higher than that of the archery group (3-per-cent level). It is interesting to note that five different instructors were involved in teaching these swimming classes and five in teaching tennis.

Little difference is found between the mean scores of the students taught by various instructors.[4] There is only one instructor whose students indicated a mean attitude which is significantly higher than that of any of the other instructors' students. This instructor taught only swimming and tennis so it is a question as to whether the activities influenced the instructor results or the instructor influenced the activity results. Since there were five instructors involved in both swimming and tennis, it seems that the activity may be a factor influencing attitude to some degree. It may be that those with the more favorable attitude toward physical education tend to elect

[2]Where groups of 20 or more are considered.
[3]Detailed figures can be obtained from the authors.
[4]Detailed results can be obtained from the authors.

these activities. As has been previously stated, it is recognized that other physical education classes previously experienced by these students will have an influence on attitudes expressed by them at this time. However, these results may indicate that it is important to sell the activity program to many of those enrolled in archery. It must be noted that the range of scores shows that scattered in most activity groups there are some students with unfavorable attitudes and, therefore, some students in every class need to be helped to realize the possible benefits of a good physical education program. Interestingly, the lowest score was made by a student who was enrolled in beginning tennis at the time. All activity groups have a wide range of scores with a negatively skewed distribution.

When the responses to the various individual items are studied on 300 score sheets, it is found that on all items except numbers 5, 12, and 24, the percentage who indicated a favorable attitude is significantly higher than the percentage who indicated an unfavorable attitude. The students were approximately equally divided in their opinions as to whether physical education is one of the most important subjects in helping to establish and maintain social standards (item 5). However, they agree with item 10 which states that physical education tends to make one a more socially desirable person. Items 12 and 24 have to do with the emphasis given physical education in the school and their responses indicate that they feel it is emphasized as it should be. It is interesting to note, however, that while 54 per cent indicated that there should not be over two one-hour periods of physical education per week (item 24), 70 per cent agree with the statement that all who are physically able will profit from an hour of physical education every day (item 33).

When the statements dealing with physical education's contribution to social growth and adjustment, to mental health, and to physical health are grouped together,[5] it is found that the percentages reacting favorably to these three groups are very similar to those reported by Bell, *et al.* (1, *p. 384*) for two groups of Freshmen at the University of Michigan (see Table 2). The student's reaction to these statements indicates that from 76 to 84 per cent feel that physical education does contribute to these three areas of development.

TABLE 2

Percentages of Students Agreeing with Statements Concerning the Social, and Physical and Mental Health Values of Physical Education

Value	University of Washington Freshmen and Sophomores	University of Michigan Freshmen	
Social	76%	73%	76%
Physical Health	84%	77%	74%
Mental Health	84%	81%	84%

[5]Items grouped as in the Bell (1, *p. 391*) study.

Conclusions

1. The great majority of the Freshman and Sophomore women students enrolled in physical education activity classes at the University of Washington expressed a very favorable attitude toward physical education as an activity course.

2. Students in swimming and tennis have a more favorable attitude toward physical education activity and those in archery react less favorably. These differences may be due to selective factors or to the influence of these activity courses on attitude.

3. There are very few significant differences between the mean attitude scores of students of various instructors.

4. While the students indicate that they felt that the time allotted to physical education activities (2 one-hour periods per week) was ample, they recognized that they would profit from an hour of physical education daily.

5. A high percentage of these students agree that physical education activity classes contribute to social development, and mental and physical health. Since this agrees with the results found in an independent study at Michigan, the results are not locally unique.

REFERENCES

1. Bell, Margaret and Etta Walters and Staff, Attitudes of women at the University of Michigan toward physical education, *Research Quarterly*, **24**: 379-391, 1953.
2. Broer, Marion R., A critical evaluation of a general basic skills curriculum for students of low motor ability at the University of Washington, *Research Quarterly*, **26**: 15-27, 1955.
3. Reeves, Warren Edgar, *A Method for Evaluating a University Physical Education Program*, Unpublished doctoral dissertation, University of Southern California, 1951.
4. Wear, Carlos L., Evaluation of attitude toward physical education as an activity course, *Research Quarterly*, **22**: 114-126, 1951.

(Submitted 12/1/54)

Objective Determination of Resistance Load for Ten Repetitions Maximum for Quadriceps Development

LT. DAVID H. CLARKE, MSC
and
LT. EDWARD L. HERMAN, MSC
Fort Sam Houston, Texas[1]

Abstract

In this study 30 subjects were given five knee extension exercise bouts based upon 30, 35, 40, 45 and 50 per cent of their quadriceps strength as measured by tests with an aircraft tensiometer. On the basis of the number of repetitions produced at each percentage, it was concluded that a resistance load equal to 50 per cent was a reasonably satisfactory method of determining the amount of weight necessary for ten repetitions maximum.

THE VALUE OF progressive resistance exercises has become generally recognized in physical medicine and rehabilitation as a technique for efficiently increasing muscular strength. A leading medical proponent of this type of exercise is Thomas L. De Lorme (6). In his method, the muscles are warmed up by lifting light loads; the final exercise bout, however, is against a resistance load which requires exertion of the muscles' maximum power for ten repetitions. A constant problem of therapists and technicians utilizing De Lorme's procedures has been to determine with as little delay as possible the maximum capacity of any muscle group, and especially to determine the resistance load for the final ten repetitions. Trial-and-error methods have largely been used, which, at best, are inadequate because valuable treatment time may be lost.

A review of the literature related to tests to measure the strength of individual muscle groups indicates limited availability of objective techniques in this area. Lovett and Martin (10) proposed a spring balance test and Elbel (7) employed the Chatillon spring scale in studying the strength of a number of muscle groups. In an investigation of various strength testing instruments reported by Clarke (2), however, the spring balance was found to be inadequate for strength testing. Therefore, a need for an instrument that would be more adequate and objective in measuring strength was seen. In this re-

[1]Lt. Clarke is Assistant Chief and Lt. Herman is Chief, Physical Reconditioning Section, Physical Medicine Service, Brooke Army Hospital, Brooke Army Medical Center, Fort Sam Houston, Texas.

search, an aircraft tensiometer was utilized to record the amount of tension the subject could apply for specified movements of the joints (1).

In the same manner, very few studies exist related to the objective determination of the amount of weight that should be used for the ten repetitions suggested by De Lorme (6). Klein and Johnson (9), in a study of quadriceps muscle development, stated that their subjects could usually produce a knee-extension exercise bout of ten repetitions maximum when the resistance load was five pounds less than the amount of weight they could lift on a single effort. However, this method is cumbersome for ease of application.

The purpose of this research, then, is to establish a practical basis for the determination of the maximum resistance load for ten repetitions for the quadriceps muscle group.

Methodology

SUBJECTS

The subjects utilized in this research were U. S. Army personnel. In all, 30 subjects were utilized, selected upon their complete willingness to co-operate at all times during the experiment. The majority of these subjects were physical reconditioning technicians and former college physical education majors. The nature of their duties and their athletic background suggested that they were physically superior to the population in general.

TESTERS

Before any data were collected for research purposes, the testers were required to demonstrate competence in the administration of knee-extension strength tests. They administered the test twice to 36 subjects and correlated the results. Objectivity coefficients, in the form of coefficients of correlation between each pair of scores obtained by different testers, were computed. The highest index of repeatability (objectivity coefficient) formerly achieved for this test was .94 (1). In their training to administer the test, the testers achieved a coefficient of .98 by this method. Thus, testing competence was considered satisfactory and use of test data for research purposes was justified.

TESTING INSTRUMENTS AND TECHNIQUES

This research is primarily concerned with the use of two testing techniques. The first is the above-mentioned cable-tension strength test which is designed to measure the strength of the knee extensor muscles, and the other is the knee extension exercise utilizing the quadriceps "boot" and weight plates. Following are descriptions of these techniques:

1. *Cable-tension strength tests.* The testing technique utilized was the cable-tension method, illustrated in Figure I, developed by Clarke and associates (4) with the following exceptions:

(a) The knee on the side tested was placed in 135 degrees extension.

(b) A pulling strap was placed around the lower leg immediately superior to the malleoli of the tibia and fibula. A stirrup-like strap was connected and passed under the instep of the foot to keep the pulling assembly in place. A ⅛-inch cable was attached to the strap, which was in turn connected to a chain and anchored to the base of the table.

The amount of tension the subject could apply to the cable was recorded with the tensiometer.

FIGURE I. Cable-Tension Strength Testing Technique.

2. *Resistive Exercises.* The method of administering quadriceps resistive exercise is illustrated in Figure II and described as follows:

(a) The position of the subject's body for each knee extension exercise was the same as for the quadriceps strength test. A pad was placed under the knee joint for better positioning and comfort.

(b) An especially made quadriceps "boot" was employed, which consisted of a leather convalescent shoe riveted to a regulation iron "boot" with an iron bar passing through the center of this at instep level. Calibrated iron weight plates, as appropriate for each subject. were placed on the bar.

(c) No warm-up bouts were permitted. The subject's exercise was confined to an alternate concentric and eccentric contraction of the extensors of the knee, with complete extension required on each movement.

(d) The subjects were instructed to continue this exercise for as many repetitions as possible; that is, until they reached the point where they failed to gain complete extension. Emphasis was placed, in particlular, upon the following two factors: (1) Lowering the weights in a smooth and even manner; (2) Reducing the underswing of the weights to a minimum.

FIGURE II. Resistive Exercise Technique.

EXPERIMENTAL PROCEDURES

Cable-tension strength tests of the knee extensor muscles were administered to both the right and left legs of the subjects at the beginning of the experiment. Five resistive exercise bouts for the same muscle group were administered on five succeeding days, the amount of weight at stated percentages of the subject's strength. The technical steps were as follows:

1. The amount of resistive weight was determined from the strength tests as percentages of the strength thus recorded. The percentages used were 30, 35, 40, 45, and 50.

2. Each bout continued until the subject had given as many complete repetitions as possible.

3. Strength tests were again administered on the day following completion of the last resistive exercise bouts.

4. In an effort to eliminate the effects of a possible increase or decrease in strength of the muscles utilized in the exercise over the five periods, and to counteract any learning factors that might be involved in this type of exercise effort, the bouts were systematically rotated in accordance with the following Latin-square arrangement:

	30%	35%	40%	45%	50%
Group 1 (nine subjects)	1[2]	3	5	2	4
Group 2 (seven subjects)	4	1	3	5	2
Group 3 (six subjects)	2	4	1	3	5
Group 4 (two subjects)	5	2	4	1	3
Group 5 (six subjects)	3	5	2	4	1

[2]Numbers indicate order of bouts for each group.

Results of the Experiment

The results of this testing appear in graphed form in Figure III, and show the following characteristics:

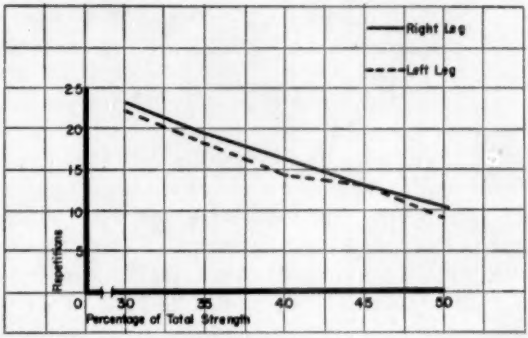

FIGURE III. Mean Repetition Curves in Quadriceps Development Exercises at Various Percentages of Strength.

1. Mean repetitions for the right leg were: 30 per cent, 23.31; 35 per cent, 19.45; 40 per cent, 16.07; 45 per cent, 13.37; and 50 per cent, 10.43. N = 30. These repetition means present nearly a straight-line decrease from the light to the heavy resistance loads.

2. The left leg exhibited nearly the same pattern of repetition decrement as did the right leg, with the following means: 30 per cent, 22.33; 35 per cent, 18.86; 40 per cent, 14.71; 45 per cent, 13.15; and 50 per cent, 9.71. N = 28. It is noticed from these mean repetitions that the decrease in repetitions was not as great at 45 per cent when compared with the right leg.

3. A comparison of mean strength scores for both legs before and after exercise sessions is shown in the following statistical computations:

Pre-exercise mean	208.45 lb.
Post-exercise mean	217.30 lb.
Difference between means	8.85 lb.
Standard error of the difference	5.62
t-value	1.57

The significance of the changes in strength as the result of the resistive exercise was tested by the null hypothesis. Using 30 as N, t-ratios of 1.70 and 2.46 denote significance at the .05 and .01 levels of confidence respectively. The standard error formula for the difference between means of correlated groups was used, inasmuch as the study was in the nature of a single group experiment (8). It is problematical whether all-out work of this type has influenced the increase in strength over the five consecutive days, or whether the learning factor has offset the strength reduction which occurs prior to improvement. It can be seen that, while strength appeared to improve, the t-value is not statistically significant.

4. As 50 per cent of the total strength best represents ten repetitions maximum (mean of 10.43 repetitions for the right leg, and slightly less for the left leg), a further analysis of the variability of the right leg scores was made. For this, the range was between five and 16 repetitions, the mean was 10.43 repetitions and the standard deviation was 2.32 repetitions. Thus, under normal conditions, approximately 68 per cent of the scores would fall between 8.11 and 12.75 repetitions.

5. Similar results were obtained with the left leg when the 50 per cent proportion of strength was used.

6. Coefficients of correlation obtained for the right leg between the number of repetitions produced at 30 per cent and those for 50 per cent, and also between repetitions at 45 per cent and 50 per cent were .38 and .55 respectively. A further analysis shows that the scatter of scores was quite different at 30 per cent than at 50 per cent. For example, the standard deviation for repetitions at 30 per cent was 5.82, as against 2.29 for 50 per cent.

Conclusions

In this study, a resistance load equal to 50 per cent of the strength of the muscles tested was found to be a reasonably satisfactory method of objectively determining the amount of weight necessary for ten repetitions maximum in exercising the knee extensor muscles. Nearly a straight line decrease in the mean number of repetitions resulted when the weights used varied from 30 to 50 per cent of knee extension strength. Similar results were obtained separately for both the right and left legs.

REFERENCES

1. Clarke, H. Harrison. *A Manual: Cable-Tension Strength Tests.* Chicopee, Mass.: Brown-Murphy Co., 1953.
2. ———, Comparison of Instruments for Recording Muscle Strength, *Research Quarterly,* **25**: 4 (Dec. 1954).
3. ———, Objective Strength Tests of Affected Muscle Groups Involved in Orthopedic Disabilities, *Research Quarterly,* **19**: 119 (May 1948).
4. ———, Relationship Between Body Position and the Application of Muscle Power to Movements of the Joints, *Archives of Physical Medicine,* **31**: 81-89(Feb. 1950).
5. Daniels, Lucile, Marian Williams, and Catherine Worthingham. *Muscle Testing.* Philadelphia: W. B. Saunders Co., 1947.
6. De Lorne, Thomas L. *Progressive Resistive Exercise.* New York: Appleton Century Crofts, Inc., 1950.
7. Elbel, Edwin R., *A Study in Short Static Strength of Muscles.* Unpublished Master's thesis, Springfield College, Springfield, Mass., 1928.
8. Garrett, Henry E. *Statistics in Psychology and Education.* 4th ed. New York: Longmans, Green and Co., 1953.
9. Klein, K. K. and E. Johnson, Research: A Method of Determining the Maximum Load for Ten Repetitions in Progressive Exercise for Quadriceps Hypertrophy, *Journal of the Association of Physical and Mental Rehabilitation.* **31**: 81-89 (Feb. 1950).
10. Lovett, R. W., and E. C. Martin, Certain Aspects of Infantile Paralysis with a Description of a Method of Muscle Testing, *Journal of the American Medical Association,* **66**: 732-733 (1916).

(Submitted 2/2/55)

Influence of Wheat Germ Oil as a Dietary Supplement in a Program of Conditioning Exercises with Middle-Aged Subjects

THOMAS KIRK CURETON
University of Illinois
Urbana, Illinois
assisted by **RICHARD H. POHNDORF**

Abstract

This study was a staff experiment carried out in 1953 as a follow-up to several preliminary experiments by graduate students at the University of Illinois, School of Physical Education. The studies were carried out in the Physical Fitness Research Laboratory. The preliminary studies as well as the staff experiment of 1953 and the reversal experiment of 1954 all showed some advantage for supplementary feedings of wheat germ oil (WGO) administered during or immediately following physical conditioning periods of conditioning exercises and swimming in an indoor pool with water temperature 74°-76° F. The subjects were adult men, 26-60 years of age, mainly sedentary, volunteers from the nonphysical education staff.

The WGO capsules (20 capsules per day, 3 minims each, and each containing 175 mg. of WGO and 0.44 mg. of mixed tocopherol) were fed for eight weeks in connection with the progressive physical training experiment. Two groups of eight men each were matched in treadmill-running time, the brachial pulse wave and age. Two other inactive control groups of 5 men each were tested at the beginning (T_1) and at the end (T_2) of the eight week period. The group results show significant advantage for the experimental subjects who took WGO over those who did not in both performance (willpower dominated) tests and in naive (non-willpower) tests.

The physiological advantage is shown in terms of running endurance in "all-out" treadmill runs, T-wave of the ECG ($CR_{IV\text{-}vlead}$), lower systolic blood pressure, the Schneider Index and the Illinois total body reaction-time test in response to light, sound, and combined signals. Individual differences in response are shown and some of the factors which cause such variations are identified.

ANNOUNCEMENT WAS first made in October 1954 of the discovery of the value of wheat germ oil (hereafter called WGO) as an ergogenic aid to men trying to improve their physical fitness in a physical training course (3, 4, 5). This first report was based upon various experiments extending from 1949 through 1952 carried out by graduate students working under the supervision of the writer. This further report is an experiment carried out in 1953 with the aid of certain staff personnel.[1]

[1]Aknowledgment is given to Richard H. Pohndorf, instructor in physical education; also to the consultants on the study: Connor Johnson, Ph.D., professor of animal nutrition; Norris L. Brookens, M.D., Ph.D., physician at the Carle Clinic, Urbana; Colin R. Blyth, associate professor of mathematical statistics; and to our staff physician who assisted with medical examinations.

The first experiment as a student thesis was carried out in 1949 by William A. Forr (17) on six groups of college students, three groups of eight in a fraternity house and three other groups of nine each in basic swimming. He found significant changes in the T-wave (highest in Lead $CR_{IV \text{ or } V}$) of the electro-cardiogram (ECG) and also in the Illinois vertical jump reaction-time test which favored WGO over the effect of vitamin E capsules (150 mg/week) and corn oil capsules of equal calorific content. The WGO was calculated to match the synthetic vitamin E (alpha tocopherol acetate). The control groups were considered to be those on the corn oil capsules. None of the subjects knew what they were taking and the administration of the capsules was carefully supervised.

Again in 1950-51, advantage was shown for a group of six men who were fed WGO capsules (20 capsules, 3 minims each, containing 175 mg. of WGO and 0.44 mg. of tocopherol) daily in connection with a strenuous physical conditioning experiment in which hard endurance training (four times a week) and a weekly "all-out" ergometer bicycle test ride was given both at ground level and at 10,000 feet of simulated altitude in a decompression chamber. After 12 weeks of preliminary training and testing, the men were given the dietary supplement. Instead of holding their position even on the plateau which they had established in the last three weeks of the training period, the men improved, four of them in a rather extraordinary manner, contrary to the usual experience of men in hard training showing evidence of terminal fatigue (staleness). Each one of the graduate students participating in this work wrote up the results on one test item. The significant results are available in the theses of Smiley (32), who showed advantages in the bicycle "all-out" ride test; Susic (35), who showed significant changes in the T-wave of the ECG; and White (37), who showed advantages in the Schneider Test. These preliminary experiments awakened considerable interest and further studies were arranged.

In 1952 three experiments were conducted. One of these by Storm (34), checked the effect which the WGO feeding had with respect to its possible tendency to fatten middle-aged men and to increase their weight. Three groups of eight men each were matched, one group was fed WGO, another was fed cottonseed oil placebos, and these were compared with the third which had no exercise or dietary supplement. While taking 20 of the 3-minim capsules per day, the WGO group showed an insignificant change in weight which averaged less than ± 2.0 pounds. This group also lost adipose tissue from 140 mm. (49.2 S.S.) to 124 mm. (58.1 S.S.), by Cureton's composite of six measurements with calipers on the cheeks, abdomen, waist, Gluteal fold, front and rear thigh. The group on cottonseed oil capsules remained about the same, 137 mm. to 138 mm. (51.5 S.S. to 50.1 S.S.). Both of these groups took the same exercise program for 12 weeks. The third group which took no exercise or supplementary feeding changed from 144 mm. (47.3 S.S.) to 125 mm. (57.7 S.S.). The WGO group closely paralleled the control group.

Two studies were done on athletic teams by Marx (28) and by Vohaska (36), using the varsity swimming and wrestling teams, respectively. The same pattern of feeding 20 capsules (3 minims) per day was used for the WGO experimental group. The experimental group was contrasted with a matched group of men on capsules of cottonseed oil containing equivalent amounts of vitamin E (2.5 mgm. of alpha tocopherol acetate per gram of cotton-seed oil) to that in the WGO. The supplementary feeding was for four weeks only. These experiments produced significant changes in several of the test variables but only the wrestling team groups were differentiated by the WGO and only in one test, namely, the brachial pulse wave. A type of fatigue called by us "competitive fatigue" developed in some of the men on both of these teams which tended to depress the results on the final series of tests, thus minimizing the changes.

In 1955, the work on the wrestling team was repeated with similar results, with trends in the right direction for improvement favoring both wheat germ oil in one group and wheat germ in another matched group, both contrasted with a control group on cotton-seed oil placebos. Some advantage showed with respect to treadmill run time, T-wave of the ECG, systolic amplitude of the brachial pulse wave and the Schneider Test. However, the same type of "competitive fatigue" showed in some of the men, supported by subjective reports of lassitude, loss of sleep, and minor injuries. Even in spite of some gains, the cardiovascular tests were lower on the final (T_2) series of tests than they would have been if the pressure had not been so great toward the end of the season. Thus, the changes obtained were minimized, and we concluded that these types of competitive sports are not very suitable to show the effects of a dietary supplement. In spite of these difficulties, the WGO was judged to be helpful.

Further research was needed and was carried on in the form of the experiment now to be described. The present experiment was carried out in 1953 with the work being done on two groups of middle-aged men with eight men in each group, and also two other groups of men classed as environmental controls. One of the latter was on the same dosage of WGO and the other took nothing; both were inactive compared to the experimental subjects. While the men competed against their own test scores, there was no competition against opponents; furthermore, the work was carefully graduated with every precaution taken to prevent the men from working too hard at first.

Experimental Design

The problem was to determine whether WGO capsules, administered as a dietary supplement for eight weeks in connection with a course in physical conditioning, would significantly differentiate the experimental WGO group from another matched group on equal size and equal calorie content placebos of devitamized lard, with both groups taking the same program of exercise.

The tests used were: (1) All-out treadmill run, (2) Brachial pulse wave (sphygmogram), (3) Illinois vertical jump reaction-time test, (4) T-wave of the ECG, (5) Standing systolic blood pressure, (6) Schneider test, and (7) Basal metabolic rate (by regular oxygen consumption procedure).

The 16 experimental subjects were tested (T_1 = Initial Test) in the last week of May and the first two weeks in June, with the tests given to each man individually by appointment. They were retested in the second week of August. The weather conditions were fairly comparable. At the final test (T_2) some of the men were busy preparing examinations to finish the summer session. If anything, the added anxiety would have hurt the cardiovascular tests. The inactive environmental controls were divided into two groups, one which took WGO and the other took no dietary supplement.

After the T_1 series of tests the 16 experimental subjects were paired man for man by age and by the time on the all-out treadmill run, just as closely as this would permit. The two resulting groups of this matching were not significantly different from each other. Some good men and some poor men were in both groups. The capsules of WGO and the devitamized lard placebos were of identical size (6 mimims). The men were not told what they were taking, except "just vitamins."

The attendance at the exercise sessions averaged four periods per week and an attendance roll was kept. The participation was about equal in terms of number of periods for the two experimental groups. In the workouts on the pool deck and in the water the men were intermixed at random. A leader set the exercises and the men did the best they could to follow them for 25 minutes, then there was a five-minute break for showers, followed by swimming for another 20-25 minutes. The pace could not always be controlled in the exercises or in the swimming, but the men kept busy at exercise to the extent that they could tolerate it. The leader was either Cureton or Pohndorf. After the first month of gradually progressive work, a series of test exercises were introduced and two or three of these were done every day after 15 to 20 minutes of warm-up work. In the sixth and seventh weeks, some similar test exercises were added in the pool. Scores were put down for all of the test exercises and it is judged that the motivation to improve was high.

No attempt was made to control the diet of any of the subjects except that all of the experimental subjects were told to eat moderately, to reduce their intake of fried foods, but otherwise to eat normally. This is all of the control that seemed wise with normal university personnel, carrying on their regular duties and home life day by day.

While the two experimental groups were fairly closely matched in age and ability, the WGO group averaging 41.4 years and 3.44 minutes for the run on the treadmill against 36.5 years and 3.34 minutes for the lard placebo group, the two groups of environmental controls who did not exercise in the program were not well matched. One of these was composed of five men, mainly

younger laboratory graduate assistants; the other of five older men who were mainly interested in just taking the tests. No attempt was made to control their eating or living habits in any way. None were in competitive sports. In the main they were quite sedentary, especially the older men.

Results

Figures I and II show the results for the treadmill run and the brachial pulse wave, respectively. The former involved exertion of voluntary willpower, whereas the latter did not involve any willpower. It is interesting that the results are much the same. This indicates that the gains[2] were not due to

---

[2]The results shown in the tables and figures are computed in terms of:

(1) Raw Score Percentages, wherein Per Cent Change $= \dfrac{T_2 - T_1}{T_1} \times 100$

(2) Standard Scores (S.S.), based upon tables from 0 to 100, equal increments of variability with 1 S.S. $= \dfrac{6 \times \sigma}{100}$

(3) Fisher's "t" tables for small samples.

(4) Standard errors of measurement from retest data:
Reliability of a single case, wherein, interpretation is limited to the subjects in the experiment and not in the universal sample,

$$s = \sqrt{\frac{d^2}{2N}}$$

d = deviation between T_{1a} and restest T_{1b}

$$t = \sqrt{\frac{T_2 - T_1}{s \vee 2N}}$$

N = number of paired retests included

or in the case of just one set of retests, T_1 and T_2, with $N = 1$; $t = D/s \vee 2$
Explanation: Assume for any given individual that the item measured is normally distributed. If X and Y are two measurement observations on the same individual, retests being taken within a few days of each other, then,
$d_1 = X_1 - Y_1$ and the mean of all such d's is 0.
variance $r^2 = 2\sigma^2$
when mean is 0, we may estimate r^2 not by $\Sigma (x_1 - x_2)^2/N\text{-}1$ with d.f. = N-1 but by the slightly better estimate $\Sigma d^2/n$ with d.f. $= n$,

therefore, $n = 1$, $\sigma^2 = r^2/2$ or $s^2 = \sqrt{\dfrac{\Sigma d^2}{2n}}$ and $s = \sqrt{\dfrac{\Sigma d^2}{2n}}$

See page 40 in Gulliksen's *Theory of Mental Tests,* 1950, for more complete derviation. This describes measurement errors and enables us to say whether differences are significant in the sense that they could not be caused by measurement errors.
Single case differences like this, then, are significant if they exceed 2.04 in absolute value, meaning that there are less than 5 chances in 100 of the differences being due to the measurement error.

Significance of the group changes were determined by $D/\sqrt{105} > 2.04$ for 5% level of confidence and applied to the two groups of 8, and > 2.75 for 1% level of confidence.

fluctuations of willpower (psychological) changes but were mainly circulatory changes and neuromuscular condition changes. The data are summarized in Tables 1 and 2, but these include additional data to the raw score data plotted in Figures I and II. The basic data are not shown at all. In Table 2 the results are shown only in terms of significance, computed case by case in relationship to the standard errors of measurement. This method shows only whether the individuals changed who were in the experiment and no generalizations about reliability of other samples can be made.

1. *Treadmill run results.* Figure I shows that the WGO actively exercised group (A) composed of eight men averaging 41.4 years of age improved from 3.44 to 5.14 minutes, or 51.5% (18.0 S.S.) on the all-out treadmill-run test. The comparable group (B) on lard placebos improved from 3.34 to 4.06 minutes, the men averaging 36.5 years of age. Both of these groups made a significant change ($t_A = 8.71$) and ($t_B = 3.63$). These two groups were also significantly different in the gains that they made ($t_{AB} = 3.59$). The five inactive controls on WGO improved from 1.90 to 2.00 minutes, an insignificant gain ($t = 1.00$); also, the five inactive controls on placebos failed to gain (2.05 minutes to 2.05 minutes). The exercised group on WGO differed significantly from the control group which remained inactive but took WGO ($t_{AC} = 4.62$); whereas the exercised group on placebos was not significantly different from the inactive control group on placebos ($t_{AD} = 1.95$). Thus, by group means the advantage of WGO stands out sharply. All changes are shown in Table 1.

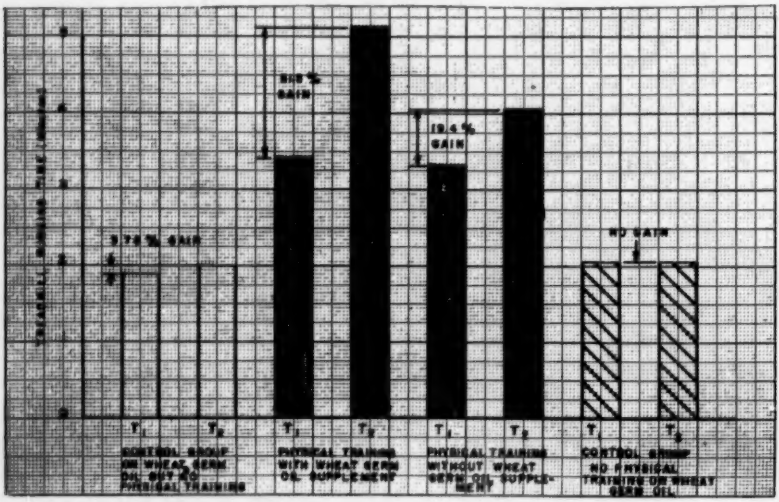
51.5 % GAIN
18.4 % GAIN
NO GAIN
PHYSICAL TRAINING WITH WHEAT GERM OIL SUPPLEMENT
PHYSICAL TRAINING WITHOUT WHEAT GERM OIL SUPPLEMENT

FIGURE I. Effect of Wheat Germ Oil and Training Upon Treadmill Running.

TABLE 1

Effect of Wheat Germ Oil Supplement on the Treadmill Run

N	Group		Avg. Age	T_1 Avg. Raw Score	S.S.	T_2 Avg. Raw Score	S.S.	D Avg. Raw Score	D S.S.	Criterion 5% Level Signif.	Obtained Group Signif.
8	A	Exercise plus WGO	41.4	3.44	51.4	5.14	69.4	1.70	18.0	2.04	8.71
8	B	Exercise plus Placebos	36.5	3.34	50.5	4.06	58.0	0.72	7.5	2.04	3.63
5	C	Inactive plus WGO	36.2	1.90	45.0	2.00	47.6	0.10	2.6	2.04	1.00
5	D	Inactive plus Placebos	32.0	2.05	47.4	2.05	48.4	0	1.0	2.04	0

Because certain individuals responded to the WGO very definitely and others did not, it was decided to make a test of each and every case, using the test of significance for an individual case, using $D/s\sqrt{2}$ for the obtained t value. The results are shown in Table 2 in terms of the number of individuals who made significant changes, those who changed in the right direction, those who did not change, and those who lost. The results again show that seven out of eight subjects in group A improved, with four making a significant change. In group B five out of eight made an improvement but only three made a significant change. This table does not show the magnitude of the changes and, of course, is not as satisfactory as Table 1. The two tables enforce the recognition that both the endurance exercise program and the wheat germ oil were significant factors in the improvements made.

TABLE 2*

Tabulation of Individual Changes (Number of Individuals)

N	Group	Signif. Gain	Gain	No Gain	Loss	Signif. Loss
8	A	4	3	0	1**	0
8	B	3	2	1	2	0
5	C	0	4	0	1	0
5	D	0	4	0	1	0

**Statistical Note, Table 2:* The s was computed from retest data for each item. The criterion of significance was $D/s\sqrt{2N}$ for each individual case, two degrees of freedom based on T_1 and T_2. *Cf.* Helen M. Walker and Joseph Lev, *Statistical Inference*, pp. 151-160, New York: Holt, 1953.

***This subject was sick at the time of the T_2 retests, but the medical report on a mild infection came too late to exclude him. It probably explains why his T_2 tests were not quite as good as his T_1 tests.

A most important test was made by running the same program in the summer of 1954 in which the men were reversed as to being on WGO or on placebos. Eight men gained 11.62 minutes while on WGO compared to the same men gaining only 1.99 minutes while on placebos; the average gain was 1.45 compared to 0.24 minutes. Out of eight reversals available, six men made their peak time on WGO, one man tied, and one man did a bit less well. The latter was in superb condition and had been in the exercise program for several years. These data are most revealing and only space prevents laying the detailed case record out in full.

2. *Brachial pulse wave results.* Figure II shows the same type of changes for the same groups but in terms of the brachial pulse wave area made on the Cameron heartograph and scaled with standardized procedures previously outlined by Cureton (7). The actively exercised group A on WGO improved from 0.359 sq. cm. to 0.550 sq. cm., a gain of 50.4% (21.3 S.S.), a significant change ($t_A = 6.63$). The other matched group on exercise and placebos (B) changed from 0.359 sq. cm. to 0.400 sq. cm., an insignificant change ($t_B = 1.51$). There was also a significant difference between groups A and B ($t = 2.49$). The control group (C), composed of five men averaging 36.2 years of age, were fed WGO over the same eight-week period and deteriorated from

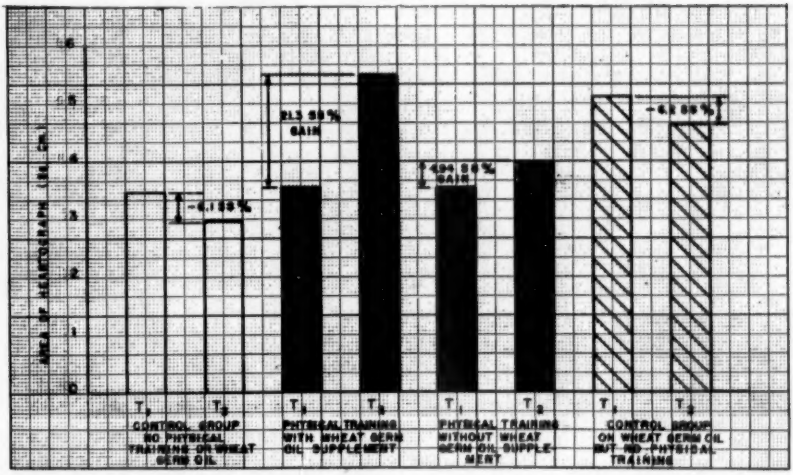
-6.1 SS %
21.5 SS %
GAIN
-6.2 SS %
CONTROL GROUP NO PHYSICAL TRAINING OR WHEAT GERM OIL
PHYSICAL TRAINING WITH WHEAT GERM OIL SUPPLEMENT
PHYSICAL TRAINING WITHOUT WHEAT GERM OIL SUPPLEMENT
CONTROL GROUP ON WHEAT GERM OIL BUT NO PHYSICAL TRAINING

FIGURE II. Effect of Wheat Germ Oil and Training Upon the Area of the Heartograph.

an average of 0.514 sq. cm. to 0.468 sq. cm., a loss of 6.2 S.S. (or 8.94% in raw score per cent). This change was insignificant ($t_C = 1.26$). The other control group (C), inactive and on placebos, also deteriorated from 0.349 to 0.299 sq. cm., a loss of 6.1 S.S. (or 14.32% in raw score per cent). This change was insignificant ($t_D = 1.62$). There was no significant difference between the exercised group on placebos (B) and the inactive group on placebos (D), ($t = 1.60$), but the exercised group improved from 0.359 to 0.400 sq. cm. while the inactive group changed from 0.349 to 0.299 sq. cm., a loss rather than a gain. These results again show the apparent value of the WGO. The results are as striking in this "naive" test as they were in the case of the all-out treadmill run.

The data from the reversal in which men on WGO were changed over to placebos and vice versa, from 1953 to 1954, both summer experiments being run under the same conditions, show that six out of twelve men improved more on WGO than they did on placebos; four men improved more on placebos; and two were the same. The average gain of the six men on WGO was 0.209 sq. cm. compared to 0.112 sq. cms. for the four men on placebos, showing that those who gain improve unusually compared to men on placebos.

The systolic amplitude of the brachial pulse wave gave similar results. The exercised WGO group gained from 1.32 to 1.76 cm., a raw score per cent gain of 33.5 per cent (16.9 S.S.). This is a significant gain at the 1-per-cent level ($t = 3.20$). The exercise group on placebos gained from 1.35 to 1.39 cm., a 2.96-per-cent gain ($t = 0.33$)—an insignificant gain. The inactive WGO control group gained from 1.41 to 1.47 cm. (4.26%); and the inactive group

of controls on placebos changed from 1.24 to 1.15 cm. (a loss of 7.24%), also insignificant ($t = 0.62$).

3. *The vertical jump reaction-time test results.* Of the three tests given in this area, the combined light and sound signal test gave the best results, this being the only test to give a significant difference. The group which exercised and took WGO gained from an average of 0.360 to 0.334, a gain of 0.026 seconds, equivalent to 8.9 S.S ($t = 2.60$). The exercised group on placebos gained from 0.336 to 0.329, or a gain of 0.007 seconds, an insignificant gain equivalent to 2.6 S.S. ($t = 0.69$). The inactive control group which took WGO changed from 0.293 to 0.295, a loss of 0.002 seconds, equivalent to 1.4 S.S. The inactive group on placebos did not change.

The visual reaction times improved a little from 0.384 to 0.358 for group A on exercise and wheat germ oil, a gain of 4.17 per cent (6.06 S.S.), but this group was not significantly different from any other group. The inactive group on WGO improved 8.50 S.S., but the gain was insignificant. The exercised group on placebos improved 1.17 S.S. and the control group on placebos improved 3.67 S.S.—both insignficant changes.

The auditory reaction time gained from 0.353 to 0.329 seconds, equivalent to 6.80% (7.9 S.S.), but the change was insignificant. The inactive group on WGO gained 5.88 S.S., and the inactive group on placebos gained 8.0 S.S. The exercised group on placebos gained 1.9 S.S., an insignificant change.

Swimming has been shown to slow these reaction times in some other experiments at the University of Illinois. It is possible that the use of swimming with these groups depressed the results.

4. *Results with the T-wave (CR_{IV-V} lead) of the ECG.* The trend of evidence favors the exercised group which took WGO, as the mean gain of this group was from 9.84 mm. to 11.04 mm., a gain of 12.20 per cent (8.3 S.S.)—a significant change ($t = 3.13$). The exercised group on placebos also made a significant change from 11.06 to 12.07 mm., a 9.13 per cent gain (5.9 S.S.) ($t = 2.80$). These two groups, A and B, were insignificantly different, so the gains are attributed wholly to exercise. The changes in the control groups and between the control groups were insignificant. Of the inactive controls on WGO, one man gained from 15.5 to 22.7 mm., a gain of 7.2 mm. and an improvement of 46.5 per cent equivalent to 87 S.S. This illustrates the type of effect which is considered extraordinary. Among seven inactive controls on placebos, only two men gained as much as a millimeter. One of these was a swimming instructor who changed from 7.2 to 9.8 mm., a control who was not as inactive as he was supposed to be, working five days a week teaching swimming. A second subject gained from 4.3 to 7.0 mm., but this subject was unusually low to begin with and the exercise improved him in an extraordinary manner. A third subject in this group of controls lost from 11.7 to 9.6 mm., probably because he was genuinely sedentary, working the entire period as a statistical clerk. Various other experiments at the University of Illinois

laboratory have shown that endurance exercise will usually improve the height of the T-wave in the C_{IV-V} lead.

5. *Systolic blood pressure (standing) results.* The group which exercised and took the WGO reduced the blood pressure from an average of 120.75 mm. Hg. to 114.5 mm. Hg., a gain of 5.18 per cent in raw scores (7.4 S.S.). This mean difference is statistically significant at the 5-per-cent level of confidence ($t = 2.25$). The blood presure rose from 118.4 to 123.0 mm. Hg. for the control group which took WGO, an insignificant change of 5.2 S.S. in the opposite direction. The changes in the active group on placebos and the inactive group on placebos were both insignificant, changes of 1.0 S.S. and 4.8 S.S., respectively.

Changes in standing diastolic blood pressure were all statistically insignificant.

6. *Results with the Schneider Index.* The group which exercised and took WGO improved from 12.75 to 15.0 in the Schneider Index. This gain of 2.25 points is equivalent to 17.65 per cent (9.7 S.S.). This was a significant change at the 5-per-cent level of confidence ($t = 2.33$). No other group made a significant change. The matched group which exercised and took placebos changed from 13.87 to 14.12, an average gain of 0.25 points; the inactive controls which took the wheat germ oil changed from 14.4 to 13.6, a loss of 0.8 points; and the inactive control group improved from 9.8 to 10.1, a gain of 0.3 points.

7. *Basal metabolic rate results.* The group which exercised and took WGO changed from −10.43 to −8.68% in basal metabolic rate, an average gain of 16.78 per cent (7.00 S.S.). This was an insignificant change. The matched group on exercise and placebos changed from an average of −18.42 per cent to −7.80 per cent, a gain of 57.6 per cent in raw score per cent and 10.62 per cent in basal metabolic rate, a change equivalent to 10.60 S.S. but insignificant ($t = 1.43$). The inactive group on WGO changed from −16.94 per cent to −13.66 per cent, or a gain of 19.37 per cent in BMR in raw score per cent units, equivalent to 3.25 S.S. The control group on placebos changed from 5.50 per cent to 0.69 per cent, a loss of 87.5 per cent in raw score per cent and 4.81 per cent in basal metabolic rate, equivalent to 5 S.S. All of these measurements were based upon the usual oxygen intake, corrected to standard conditions, using a Sanborn metabolator. Relatively, these changes tend to show that there is a trend for exercise to raise the metabolic rate.

Discussion of Confirmation Work

The gains in treadmill-run times seem to indicate real gains in favor of WGO. When these gains were obtained in connection with the preliminary experiments they were thought to be due to the influence of training, but as various experiments have been conducted of diverse pattern and some of the experiments repeated, it is certain that some of the gain should be accredited

to WGO. The support for this is not only in the advantages shown in the matched group experiment of 1953, but also in the 1954 follow-up experiment in which the men were reversed. The fact that the gains are significant in will-power performances like the treadmill run and reaction-time vertical jump and that they were also significant in the naive tests like the brachial pulse wave area, brachial pulse wave systolic amplitude, T-wave of the ECG and standing systolic blood pressure would certainly seem to be a strong case.

When we had obtained such favorable evidence on human subjects in 1953 and 1954, the president of the Viobin Corporation, Mr. Ezra Levin, arranged to have the experiment checked by Dr. Benjamin Ershoff, of the Department of Biochemistry and Nutrition, at the University of Southern California. He chose to simulate and repeat the work on guinea pigs, swum to exhaustion with weights tied to their legs in a barrel of water. Each one was carefully timed. In groups and on several contrasting diets, the results always favored the WGO diet. Since this work was published after our paper was written and reported to the American Physiological Society, the reference to Evshoff and Levin's work is given below in a footnote.[3]

In view of the studies already reported on the value of the T-wave (CR_{IV-V} lead) to predict endurance (3, 5, 9, 29, 38) and the similar work on the value of the brachial pulse wave as standardized by Cureton (7, 12, 37), it is logical to expect the improvement to show in these tests if the gain appeared in the treadmill run. The previous work showed good relationships of the naive tests to the all-out treadmill-running performance. Then it was not much of a surprise to find all of these relationships going together in this present work. It is plain to us that many debilitated subjects have T_{IV-V} waves as low as 2 to 3 mm. and most of these can hardly walk on the

---

[3]B. H. Ershoff, and E. Levin, Beneficial Effect of an Unidentified Factor in Wheat Germ Oil on the Swimming Performance of Guinea Pigs, *Federation Proceedings*, (Am. Institute of Nutrition) **14**: 431-32, Mar. 1955.

Male guinea pigs ranging from 250 to 300 gm in body wt. were fed ad libitum highly purified rations containing all known nutrients but differing in the amount and source of dietary fat. Corn oil or wheat germ oil were incorporated in the diet at levels of 2%, 5%, or 10% of the ration. A control group was fed a natural food ration (Purina Rabbit Pellets) supplemented daily with ascorbic acid. After 28 days of feeding, swimming tests were conducted on all guinea pigs. A weight equal to 2.5% of the animal's body weight was attached to each of the hind legs. Animals were placed in a barrel, filled to a depth of 18 inches with water. The water temperature was maintained at 37° C. All the animals fed the natural food ration drowned within 10 minutes. From 25% to 33.3% of the guinea pigs fed the corn oil rations swam for 60 minutes, at which time the swimming tests were terminated. In contrast to the above groups, over 60% of the guinea pigs fed the wheat-germ-oil-containing rations were still swimming after 60 minutes. Diets containing 2% wheat germ oil were as effective as rations containing 10% wheat germ oil in promoting swimming performance. Increasing the tocopherol content of the corn-oil-containing rations by the addition of 1 gm of a tocopherol acetate/kg of diet was without significant effect on swimming performance. In contrast to findings in the guinea pig, weanling rats fed purified rations containing wheat germ oil for 28 days did not swim longer than these fed diets containing corn oil.

treadmill; the average of the University of Illinois students is 9.5 mm.; the Olympic swimming and diving men averaged 17.0 mm.; a collection of top track and field stars averaged 15.9 mm.; John Marshall measured 18.7 mm. in our laboratory and shortly afterward broke five world's records in swimming races from 200 to 600 yards; then finally Roger Bannister, first to break the four-minute mile standard, was predicted to accomplish this feat after our work on him in the University of London's Royal Free College of Medicine, in 1952, when we found his T-wave to be 27.5 mm. Many of our other experiments have shown the gain in this important measurement as associated with improvements in endurance.

It is illogical to attribute the gains in endurance to synthetic vitamin E or to the natural vitamin E in WGO. Since each of the 3-minim capsules contained only 0.44 mgm. of mixed tocopherols, at the most 21.1 mgm. could have been accounted for, whereas the normal diet contains from 30 to 100 mgm. per day. The results point to some unidentified factor in wheat germ oil which is also partially contained in corn oil. So far, direct biochemical assays have failed to find any hormones in the oil which is classed wholly as a food. One report indicates that the WGO produces "hormone-like" effects in which the combs of chickens grow faster and the gonads increase in weight (27).

It is interesting to see that some subjects are helped so much in endurance or in tests which are correlated with endurance, whereas others are very little affected. This type of variation has been reported by Embden, Grafe, and Schmitz (15) in their study of the effect of phosphoric acid upon muscular work capacity of their subjects, wherein subject *Schr.* made a clear-cut increase, whereas subject *Hoe.* did not. Such individual variations have also been shown by Droese (11) in his study of the effect of dextrose plus vitamin B_1, wherein it is explained that persons whose vitamin B_1 stores are adequate to handle extra sugar at high temperature differ from those whose vitamin stores are inadequate to handle the same sugar. In our own work in Susic's (35) experiment in which men were trained for ten weeks on a hard endurance-type program, one subject, W. S., of slight body build improved from 13.35 to 13.85 mm. in the ECG T-wave (CR_{IV-V} lead) but after administration of WGO for the last six weeks of the experiment raised this measure to 22.65 mm. This gain is not only contrary to what occurs in hard training as a number of our subjects have shown staleness and lost amplitude in that state, but the amount of the gain is not paralleled in ordinary training experiments to our knowledge. In this same experiment, subject H.C., of excellent athletic build, changed from 9.00 mm. in the ten weeks of preliminary training, to 10.75 mm., then after six weeks of taking the WGO measured 10.35 mm., showing practically no effect. Such individual differences force the study of each and every case.

It is clear in our work that the men who began their training in the experiment at a high level of physical fitness, brought about by previous training, did not improve as much in the continued training. A man like our subject *Rae* even lost a bit, probably owing to over-training. These individual variations complicate the experiment. It is impossible to know the exact level of need or tolerance for WGO and to start all men from the same relative level. So long as we take a random sample of men from all walks of life great differences will be found in the way they respond. In the light of this, the group mean differences are seen to be minimized results. Any average which includes some men who cannot respond will lower the level of response for some who had a great need or who were able to utilize and respond to the effects of WGO. We came to the computation of significance for change in each and every subject.

Table 2 shows an example of the results of such work but even this table does not show what happened in detail to every man. The basic data are not published but they are on file in our department—data which show the changes for each man in detail. These are the data which seem to us to be most valuable. We have no better way than trying each man to see if he will respond to the WGO after a period of training long enough to have the plateau established. While changes of significance are shown for the group data in treadmill-run time, brachial pulse wave, combined light and sound vertical jump reaction time, T-wave of the ECG, systolic blood pressure, and Schneider Index, some individuals are strongly affected while others are not. Trends in the right direction for improvement are shown in basal metabolism (to raise it), in reaction to light and sound signals in the total body reaction time test, and in several of the combination cardiovascular tests like the Barach Index. Several of these latter items are not reported in full.

Two other effects seem of importance. One of these is to note the effect of wheat germ or WGO on a group of men who have been relatively confined and who are not in good condition for endurance or cardiovascular condition. One such group of workers were studied who were employed in an automobile plant. We used them as a group of controls. It was a mistake not to match them carefully with our experimental groups because they made significant improvements in heartograph area and amplitude, all of the reaction times, and in some of the other quiet cardiovascular tests when put on WGO or wheat germ supplementation, but they did not improve in the treadmill run. The nutritive supplements appear to be helpful to the cardiovascular measures and neurological speed but these phases must be studied more carefully.

Workers in physical fitness (physical education, performance physiology, and sports medicine) have long striven to discover food supplements which will increase physical performance. Long documentary reviews of this line of research have been provided by Hellebrandt and Karpovich (18), P. V. Karpovich (23), Austin Henschel (20, 21), and Ancel Keys (25). While a

few studies using performance criteria have been reported in the literature (1, 2, 10, 11, 14, 15, 16, 22, 24, 26, 30, 31), we do not seem to find a study which has shown the effects in a parallel way in both the willpower exertions (treadmill run) and te naive tests (brachial pulse wave and T-wave). It seems very important to have both sets of tests in any such experiment.

Few studies have produced favorable evidence for improved physical performance which has remained uncontradicted, holding in mind the studies of vitamin B_1 (10, 11, 20, 21, 22, 23), B-complex (14), B_{12} (30), the gelatin studies (24, 27), alcohol (1), caffeine (16), the basic (alkaline) diets (10), and ascorbic acid (30). However, a fault seems to be present in most of these feeding experiments, namely, too short a time for the dietary supplement to act. In some studies, only two or three days have been allowed (23); in others, as long as a week (23); and only five to six weeks in important U. S. Army ration studies (26). We also believe that the mixing of WGO with other types of food in the stomach may minimize its utilization or destroy some of its value. To feed it on an empty stomach just before or just after exercise seems to increase its effect as borne out in our own experiments.

Conclusions

The following conclusions have been drawn for the work reported:

1. Wheat germ oil (WGO) was found to be a valuable food supplement which helped the endurance of middle-aged men to run the all-out treadmill-run test, and produced significant gains over a matched group which took the same course of conditioning exercises for eight weeks.

2. The significant gains in running endurance were paralleled by significant gains in non-willpower (naive) tests, such as: the brachial pulse wave of the upper arm (heartograph area and systolic amplitude), the T-wave (CR_{IV-V} lead) of the ECG, systolic blood pressure (to lower it) and the Schneider Index. These appear to be important cardiovascular effects.

3. Significant gains have been found in the speed reaction of the total body (using the Illinois total body reaction time test), and near significant findings on light and sound reactions which are in the right direction and favor the groups which have taken WGO.

4. Trends for the basal metabolic rate to rise (oxygen consumption tests) are noted which are more attributable to the type of exercise taken as shown in other experiments in our series. WGO did not depress the BMR as vitamin E is known to do but produced a slight rise. The group on exercise and swimming with placebos of devitaminized lard raised the BMR about 11 per cent in per cent BMR units, equivalent to 57.6 raw score per cent, whereas the control groups changed very slightly or not at all.

5. Individual sensitivity to WGO enforces the recognition of studying the reactions case by case rather than wholly in terms of group mean changes. These computations, using standards for significance for each individual

change, in relation to the standard errors of measurement ($s\sqrt{2}$), show significant gains for the subjects on WGO in all-out treadmill running as well as in the non-performance tests of cardiovascular condition.

REFERENCES

1. Asmussen, E. and O. Böje, The Effect of Alcohol and Some Drugs on the Capacity for Work, *Acta. Physiolog. Scand.*, **15**: 109, 1948.
2. Barborka, B. J., E. E. Foltz, and A. C. Ivy, Relationship Between Vitamin B-Complex Intake and Work Output in Trained Subjects, *Jour. of the A.M.A.* **122**: 717-720 (July 10, 1943).
3. Cureton, T. K., Effect of Wheat Germ Oil and Vitamin E on Normal Human Subjects in Physical Training Programs, *Am. Jour. Physiol.* (abstract) **179**: 628, (Dec. 1954).
4. ———, Exercise, Wheat Germ Oil for Middle Aged Men, *Science News Letter*, **66**: 216 (Oct. 2, 1954).
5. ———, Wheat Germ Oil, the Wonder Fuel, *Scholastic Coach*, **24**: 36-37, 67-68 (Mar. 1955).
6. ———, *Physical Fitness of Champion Athletes*, pp. 314-350: Treadmill Tests of Maximal Physical Efficiency. Urbana: University of Illinois Press, 1951.
7. ———, Ibid., pp. 228-254: The Brachial Pulse Wave of Cardiovascular Condition.
8. ———, Ibid., pp. 94-102: Vertical Jump Reaction Times of Champion Athletes.
9. ———, Review of Research to Determine Cardiovascular Condition (1941-1950), *Proc. of the Am. Assn. for Health, Physical Education and Recreation* (56th National Convention, Detroit, Research Section), pp. 167-177, Washington, D. C. The Association: 1201 Sixteenth St., N.W., 1951.
10. ———, Diet Related to Success in Competitive Swimming, *Beach and Pool*, **7**: 335 (Nov. 1933); also, **8**: 10 (Jan. 1934).
11. Droese, W., Die Hebung der Leistungfahigkeit Durch B_1 und die Erkennung Einer B_1-Hypovitaminose, *Arbeitsphysiologie*, **11**: 338-360, 1941.
12. Dupain, G. Z., Specific Diets and Athletic Performance, *Research Quarterly*, **10**: 33-40 (Dec. 1939).
13. Editorial: "Care and Feeding of Athletes," *Newsweek*, p. 102, (Mar. 23, 1953).
14. Egana, E., R. E. Johnson, E. Bloomfield, L. Brouha, A. P. Meikeljohn, J. Tittenberger, R. C. Darling, C. Heath, A. Graybiel, and F. Consolazio, The Effects of Diet Deficient in the Vitamin B-complex on Sedentary Men, *Am. Jour. Physiol.*, **137**: 731, 1942.
15. Embden, Gustav, Edward Grafe, and Ernst Schmitz, About the Increase in Productivity Through Phosphate Supply, *Zeitschrift für Physiologische Chemie*, **113**: 67-107, 1921.
16. Foltz, E. E., A. C. Ivy, and B. J. Barborka, The Use of Double Work Periods in the Study of Fatigue and the Influence of Caffeine on Recovery, *Am. Jour. of Physiol.*, **136**: 79-85, Mar. 1942.
17. Forr, W. A., *The Effect of Wheat Germ Oil and Vitamin E on Physical Fitness*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1950. 92 pp.
18. Hellebrandt, F. A., and P. V. Karpovich, Fitness, Fatigue and Recuperation, *War Medicine*, **1**: 745-768 (Nov. 1941).
19. Henry, F. M., Influence of Athletic Training on the Resting Cardiovascular System, *Research Quarterly*, **25**: 28-41 (Mar. 1954).
20. Henschel, Austin, Diet and Muscular Fatigue, *Research Quarterly*, **8**: 280-285 (Oct. 1942).
21. ———, Vitamins and Physical Performance, *The Journal-Lancet* (Minneapolis), **63**: 355-357 (Nov. 1943).

22. KARPOVICH, P. V., and N. MILLMAN, "Vitamin B_1 and Physical Endurance," *The New England Journal of Medicine*, **226**: 881-882, (May 28, 1942).
23. KARPOVICH, P. V., Ergogenic Aids in Work and Sport, *Supplement to the Research Quarterly*, **12**: 432-450 (May 1941).
24. KACZMAREK, R. M., Effect of Gelatin on the Work Output of Male Athletes, Non-Athletes and Girl Subjects, *Research Quarterly*, **11**: 109-119, (Dec. 1940).
25. KEYS, ANCEL, Physical Performance in Relation to Diet, *Federation Proceedings*, **2**: 164-187 (Sept. 1943).
26. ——— and A. F. HENSCHEL, Vitamin Supplementation of U. S. Army Rations in Relation to Fatigue and the Ability To Do Muscular Work, *Journal of Nutrition*, **23**: 259-269 (Mar. 1942).
27. LEVIN, EZRA, JOHN F. BURNS, and V. K. COLLINS, Estrogenic, Androgenic and Gonandotrophic Activity in Wheat Germ Oil, *Endocrincology*, **49**: 289-301, (Sept. 1951).
28. MARX, ELIZER I., *The Effect of a Dietary Supplement on Varsity Swimmers*, Unpublished M. S. thesis, Physical Education, University of Illinois 1952. 40 pp.
29. MASSEY, B. H., *Prediction of All-Out Treadmill Running from Electrocardiograph Measurements*, Unpublished M.S. thesis, Physical Education, University of Illinois, 1947. 102 pp.
30. MONTOYE, H. J., P. J. SPATA, VIRGIL PINCKNEY, and L. BROWN, Effects of Vitamin B_{12} Supplementation on Physical Fitness and Growth of Young Boys, *Journal of Applied Physiology*, **7**: 589-592 (May 1955).
31. ROBINSON, S., and P. M. HARMON, Effects of Training and of Gelatin on Muscular Work, *Am. Jour. Physiol.*, **133**: 161-169 (May 1941).
32. SMILEY, WILLIAM A., *Variation in a Bicycle Ergometer Test Due to Physical Training and a Dietary Supplement*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1951. 91 pp.
33. STATON, WESLEY M., The Influence of Ascorbic Acid in Minimizing Post-Exercise Muscle Soreness in Young Men, *Research Quarterly*, **23**: 356-360 (Oct. 1952).
34. STORM, WALTER N., *The Effects of Training and a Dietary Supplement on Muscle Symmetry and Fat Distribution of Adult Males*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1952. 43 pp.
35. SUSIC, S., *Effect of Training and a Dietary Supplement on the T-wave of the Electrocardiogram at Ground Level and at 10,000 Feet Simulated Altitude*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1953. 55 pp.
36. VOHASKA, WILLIAM J., *The Effects of Wheat Germ Oil on the Cardiovascular Fitness of Varsity Wrestlers*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1952. 43 pp.
37. WHITE, C. H., *The Effect of Physical Training and a Dietary Supplement on the Schneider Index at Ground Level and at 10,000 Feet Simulated Altitude*, Unpublished M.S. thesis, Physical Education, University of Illinois, 1951. 96 pp.
38. WILLET, ALBERT E., *The Prediction of Treadmill Running from Heartometer Measurements*, Unpublished M. S. thesis, Physical Education, University of Illinois, 1948. 73 pp.
39. WOLF, J. GROVE, Effects of Posture and Muscular Exercise on the Electrocardiogram, *Research Quarterly*, **24**: 475-490 (Dec. 1953).

(Submitted 11/1/54)

Instruments To Measure Strength

PAUL A. HUNSICKER AND RICHARD J. DONNELLY[1]
University of Michigan
Ann Arbor, Michigan

Abstract

The various devices which have been used to measure human strength during the past 250 years are illustrated and their principle of operation described. An effort was made to select instruments which were unique when first employed or which were radical modifications of original devices. Only the early investigators are mentioned in connection with any single instrument.

WHILE MAN HAS undoubtedly competed with his rival in feats of strength since pre-historic times, it was not until the modern historical era that attempts were made to measure a man's strength scientifically. The first study of this kind was conducted by the French scientist, De La Hire, who published the results of his investigation in 1699 (11). De La Hire did not use a highly scientific device, but merely compared the strength of men in lifting weights and carrying burdens with that of horses. Since that day, however, many different pieces of apparatus have been designed to measure the muscular strength of humans. The purpose of this article is to give a historical picture of the development of these various instruments and the principle of operation upon which their design is based.

Early Dynamometers

The first person to use an instrument which was given the name of dynamometer was an Englishman named Graham. It remained, however, for Desaguliers, a Frenchman who was a professor at the University of Oxford in England, to attract attention to Graham's dynamometer. He adapted this device to fit his own needs. Basically, he developed a machine which consisted of "a large wooden frame and stout uprights to offer points of support or resistance, and handles attached to a crossbar which moved a steelyard. By hanging weights on the latter the force used could be directly determined." (5, *p. 544*).

The next development in dynamometers came from Leroy, a member of the French Academy of Science. His instrument "consisted simply of a metallic tube, within which was placed a spiral spring with an attached graduated rod terminating above in a globe. This was to be grasped by the hand, and the spring compressed from above, the amount of force exerted being marked on the rod." (5, *p.544*).

Spring Steel Dynamometers

The first eliptical spring dynamometer of metal was produced in 1807 by a Frenchman named Régnier (1, *p.277*) and is shown in Figure I. This is really the first practical dynamometer and the forerunner of most spring

[1]Dr. Donnelly is now at University of Minnesota.

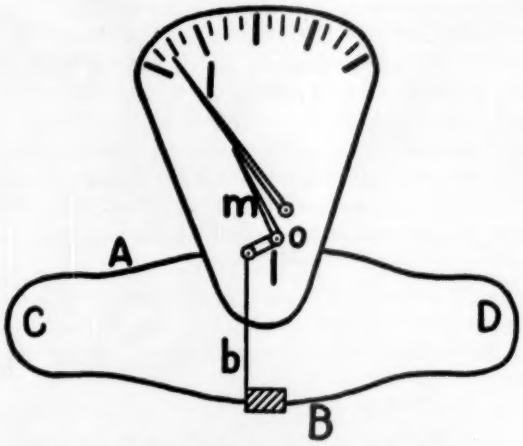

FIGURE I. Régnier Spring Steel Dynamometer.

steel dynamometers in use today. It was used to measure grip, pulling power of the arms, and lifting power of the back muscles.

Two researchers, Binet and Vaschide (3, *p.245*), have questioned the fact that Régnier invented the dynamometer bearing his name, and have suggested that a student of Régnier, named Harneiter, should be credited with inventing the elliptical spring dynamometer. However, Broca, from whose writings Binet and Vaschide obtained their information, clearly states that Régnier had already worked with an elliptical spring dynamometer before Harneiter constructed his instrument (4, *p.45*). The confusion evidently arises because the name of Régnier was stamped on the dynamometer constructed by Harneiter.

In 1859 Burq, also a Frenchman, invented another modified spring steel dynamometer (6). His dynamometer had the spring steel protected by a rectangular ring of metal. The dial was placed inside the spring steel to which two permanent handles were attached for testing compression (grip). Traction or pull could be measured by inserting handles with hooks into the ends of the permanent handles.

In 1863 Duchenne developed a dynamometer patterned after the same principle as Burq's dynamometer (24, *p.9*). Duchenne's dial was placed differently, and his handles were slightly modified. Duchenne advocated testing muscles on basically the same principle as later proposed by Martin (22). For example, to test the flexion of the forearm the subject grasped the leather strap with his forearm flexed. The tester then held the other handle and pulled on it until the subject's forearm began to extend. This was what Martin termed the "Break" method of muscle testing.

Francis Galton, as Chairman of the Anthropometric Committee of the British Association for the Advancement of Science, conducted a series of

anthropometric measurements on adult males in the British Isles over a period of six years starting in 1875 (13). One of the test items was the strength of arm taken with Galton's dynamometer which was an ordinary spring scale with a circular clock-like face. At either end of the scale was a loop through which a strap was threaded. The straps in turn were each attached to a small round wooden bar. This test of arm strength was executed by grasping one of the wooden bars with one hand and extending that arm straight out from the side as nearly as possible in the line of the shoulders. With the other hand the testee grasped the other wooden handle and tried to draw it back towards his ear as an archer would in drawing a bow.

Toward the close of the 19th century, two dynamometers were widely used in France, one invented by Mathieu and the other by Collin (see Figure II). Both were similar in design, being modifications of Régnier's dynamometer. Both were used in the United States in the latter half of the 19th century (5, 8, 17). An American physician, Hammond, objected to Mathieu's dynamometer. He disliked the fact that the subject's hand covered the dial when squeezing Mathieu's instrument (16). Consequently, he placed his dial outside the spring steel at one extremity of the ellipse.

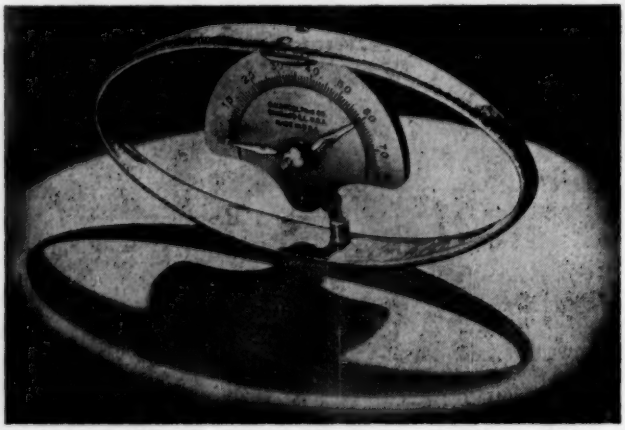

FIGURE II. Collin Elliptical Spring Steel Dynamometer.

The early developments in dynamometers in the United States are closely associated with Sargent's work in testing Harvard University students in the last quarter of the 19th century. So far as can be determined, Sargent became acquainted with dynamometers through an anthropologist, Brigham, whom Sargent knew in his first year at Harvard, 1879 (2, *p.37*). Brigham had done strength testing as early as 1872 with a case of dynamometers which he had obtained from Paris. He showed these to Sargent who immediately

ordered a set from Paris. Sargent initiated strength testing with dynamometers at Harvard in January, 1880 (26). He used one type of dynamometer similar to Régnier's instrument for measuring back and leg strength (see Figure III) and another kind for testing strength of grip. This grip dynamometer was small and compact enough to fit inside a person's closed hand. Sargent also utilized an adapter with this grip dynamometer in order to test arm-pushing and arm-pulling strength.

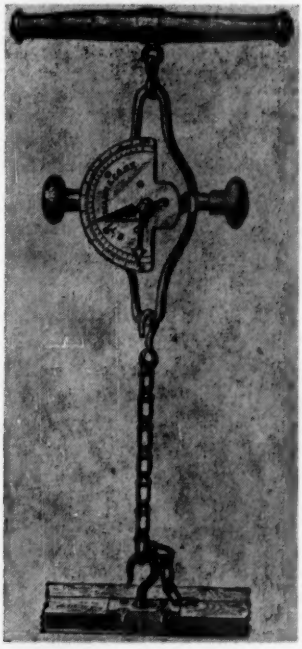

FIGURE III. Sargent Back and Leg Dynamometer.

Just prior to the turn of the present century an American anthropologist, Smedley, found that the ordinary spring steel dynamometers then in use were entirely unsuited to the size of young children's hands (28). This led to the development of the Smedley dynamometer which permitted adjustment of the grip device to fit the subject's hand (see Figure IV). There is occasional reference to this in the literature as the Stoelting dynamometer, taken from the name of the company which manufactures it.

While a physiologist in the Harvard Medical School, Martin became interested in strength testing as a result of the polio epidemic that swept Vermont in 1914. He needed some instrument by which he could evaluate

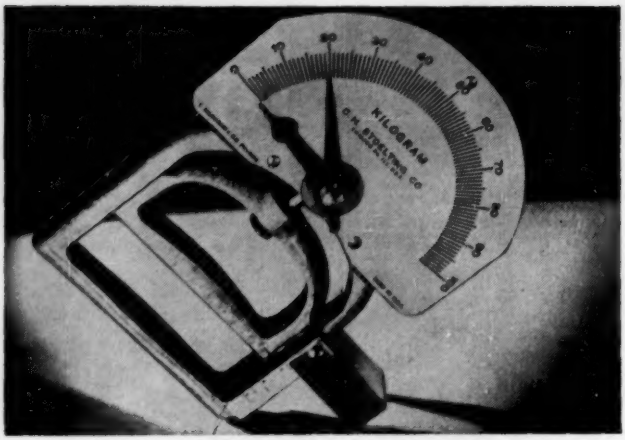

FIGURE IV. Smedley Adjustable Grip Hand Dynamometer.

the effects of his treatment. Martin departed quite radically from the previous method of testing muscular strength. Instead of using the utmost positive muscular effort that a person could elicit, he measured the "breaking strength" of muscles, *i.e.*, the muscular force required to overcome maximal resistance (22). An ordinary spring balance scale served as his dynamometer. One end of the scale was attached to a strap which in turn was fastened to the subject, while the other end of the spring balance scale was held by the tester or operator. For example, to test the biceps, the strap was fastened to the forearm of the subject who flexed his arm and offered maximal resistance to the tester's opposite pull at the command "Hold Back." As soon as the arm started to extend the breaking point occurred.

In the early 1920's Vernon did extensive work in testing workers in industrial plants in England under the sponsorship of the British Government's Industrial Research Board. He used a back and leg dynamometer similar to the one employed by Sargent. The main difference was that the chain passed over a ballbearing pulley at the base of the dynamometer, permitting the individual to pull and lift from a number of positions (32).

More recently, Schmier and his colleagues, aided by a grant from the National Foundation for Infantile Paralysis, Inc., developed an improved spring balance dynamometer patterned after the principle of Martin's. This improved form of dynamometer removed the operator from playing an active part in the test, thus eliminating one of the severest criticisms of Martin's technique and improving the validity of the test(27).

Mercurial Dynamometer

In the final decade of the 19th century, Kellogg originated an elaborate dynamometer for testing 22 muscles and muscle groups (20). He contended

that the usual methods of anthropometry were not satisfactory for measuring the results of a prescribed program of exercise designed to improve impaired muscles of invalid adults. He desired a dynamometer that would test more muscles. With his instrument, Kellogg tested several thousand invalid adults who were confined to the Battle Creek Sanatarium in Battle Creek, Michigan.

Kellogg's dynamometer is shown in a side view in Figure V. It consists of a cistern filled with mercury that is mounted on a carriage that can be adjusted according to the height of the subject by sliding up and down two vertical rods. Inside the cistern is filled with mercury and a loose-fitting cylindric float is positioned on top of it. The mercurial dynamometer operates on the hydraulic principle. A system of levers that connects the grip handles to the nipple on top of the float makes it possible to have a smaller cistern

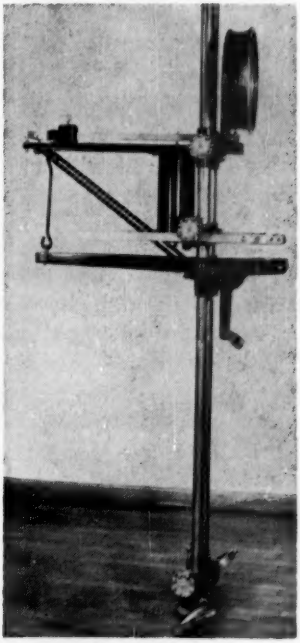

FIGURE V. Kellogg Mercurial Dynamometer.

than was possible in earlier models of this dynamometer. As the mechanical lever system is actuated by the subject, the float is forced downward into the mercury. The float is mechanically connected to the circular dial recording device. The distance the float is depressed into the mercury is reflected in the dial reading.

At the time when the spring steel dynamometers were being introduced into the United States, an American physician, named Hamilton, raised objections to the dynamometer of Mathieu. In his therapeutical work Hamilton desired an instrument that would adequately test the strength of grip in certain patients who had extremely weak hands (15). He devised a pneumatic type of dynamometer. It consisted of a glass tube closed at the top and a rubber bulb filled with colored water at the bottom. To test the grip, the subject squeezed the rubber bulb, causing the water to rise and compressing the air in the glass tube. A graduated scale marked in pounds on one side of the tube gave the reading.

In 1939 Geckler developed a pneumatic dynamometer, illustrated in Figure VI, that consisted of a rubber bulb connected by means of a short tube to an air compressor gauge (14). The amount of force as the bulb is squeezed is transmitted to the air gauge and the reading is taken directly.

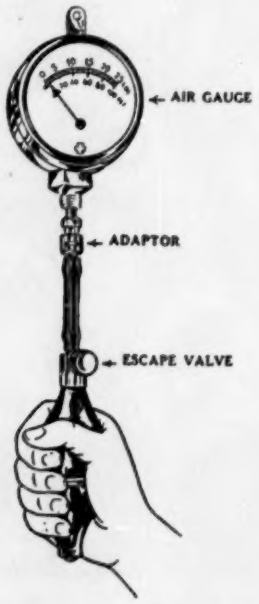

FIGURE VI. Geckler Pneumatic Dynamometer.

The Ergograph

The term "ergograph" was coined and first defined in an article by Mosso which appeared in 1890 (23). Figure VII is a reproduction of the instrument described in this paper. Mosso and his co-workers had been conducting

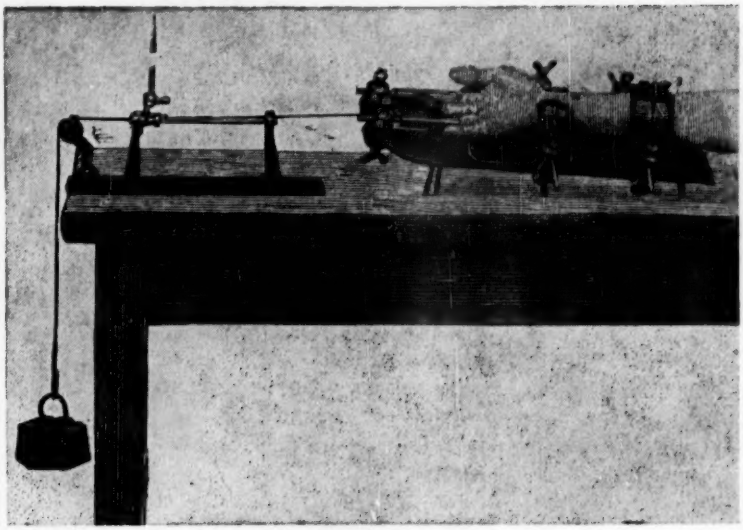

FIGURE VII. Mosso Ergograph.

numerous fatigue studies with this device for the previous four years. This differed from earlier strength testing instruments because it recorded a series of repetitious movements rather than a single maximum effort. As a result of the development of the ergograph, physiologists were able to study the many facets of muscular fatigue. During the next half century a number of ergographs were developed by a host of workers and the instrument became standard equipment in physiology laboratories.

A recent modification of the ergograph by Hellenbrandt and Kelso (18), Figure VIII, is particularly useful for orthopedic cases where a limited range of movement may occur. However, Mosso's principle of lifting a known weight a certain height in a defined time-sequence is honored. The Hellebrandt-Kelso ergograph does incorporate a host of mechanical advantages; namely, no overhanging parts, an electromagnetic signal for indicating the number of contractions, the load moves on two vertical rods with a minimum of friction, and the cumulative height of successive lifts is automatically indicated by a distance meter.

Early experiments with the ergograph provided quantitative evidence that the strength involved in this type of action was not necessarily related to the strength called for in a single maximum effort. Later workers coined terms like "strength-endurance" or "dynamic strength" for the former and "static strength" for the latter.

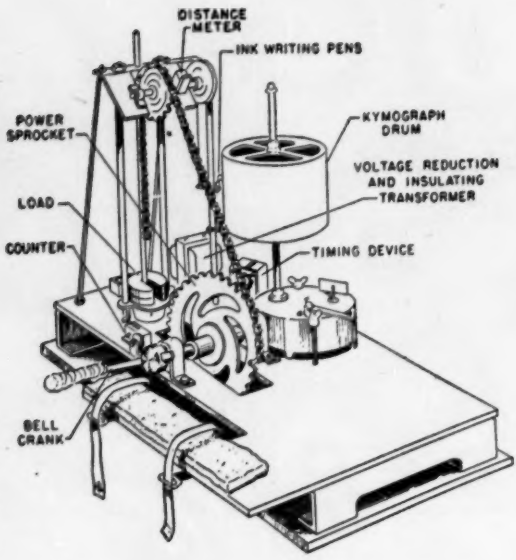

FIGURE VIII. Hellebrandt and Kelso Ergograph.

The Tensiometer

Clarke (9) adapted an aircraft control cable tensiometer for strength testing. Cable tension is determined by measuring the force needed to create offset on Riser (A) in the cable between Sector (B') and Sector (B), Figure IX. By using suitable strap and pulley attachments, Clarke was able to test the strength in various movements at the wrist, elbow, shoulder, hip, knee, and ankle joints. Because the instrument is small (4 in. x 4 in. x 1¼ in.) it is particularly useful in a test battery that involves testing the strength of movements in a number of joints. In a study comparing objectivity coefficients, Clarke found that the cable tensiometer had greater precision than either a strain gauge, spring scale, or myometer (10).

Strain Gauge Dynamometers

The most recent advance in the field of strength testing is based upon the principle underlying the electrical strain gauge. Although strain gauges have only been used for about 20 years, the theory of a wire changing its electrical resistance when stretched was advanced in Thomson's classic experiment about 100 years ago (29). Present-day gauges consist of a wire of 0.001 in. diameter fixed in a circular steel ring or an L-shaped bar. Any deformation of the bar or the ring will change the electrical resistance of the strain gauge.

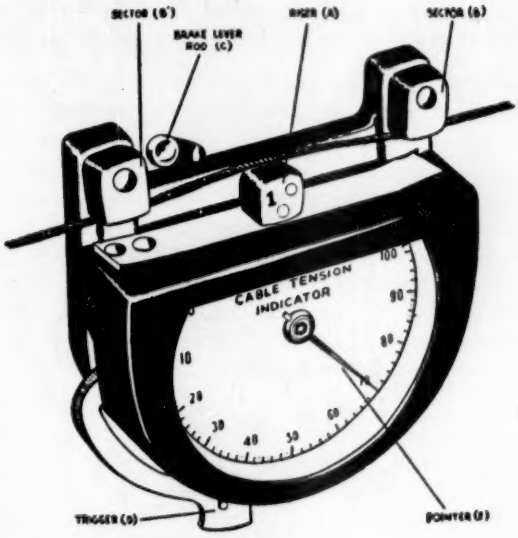

FIGURE IX. Tensiometer.

By use of the appropriate amplification and recording system, the amount of force applied can be determined.

In the past decade physiologists adopted the strain gauge. Ralston, *et al.* (25) made a study of "Mechanics of Human Isolated Voluntary Muscle," using calibrated aluminum rings to which were affixed SR-4 type A-1 electrical strain gauges. Another study reports the use of an apparatus for eliciting a stretch reflex and recording electrically the isometric responses of the muscle, by means of strain gauges, together with the summated action potentials in the muscle or nerve supplying it (7). Tuttle reported using strain gauges in a study of maximum grip strength and grip strength endurance (30). The same principle of strength testing was employed in a later study on back and leg strength by Tuttle and co-workers (31).

The anthropology Section of the Aero-Medical Laboratary of Wright-Patterson Air Force Base recently had a Kinematic Muscle Study Machine built which employed strain gauges and could test forces in three directions simultaneously (Figure X). One of the authors of the present paper had the opportunity of supervising a project employing this apparatus (19). The handle of the dynamometer is constructed around a swivel-joint and a force can be exerted in any one or all of three planes. The strain gauges are arranged to pick up distortion in the various planes. Since each axis has a separate amplifier and recording dial, it is possible to obtain data on the amount of force that is being exerted in each of the three planes simul-

taneously. The instrument is ideally suited for fatigue studies, because the total strength effort as well as the maximum value can be recorded.

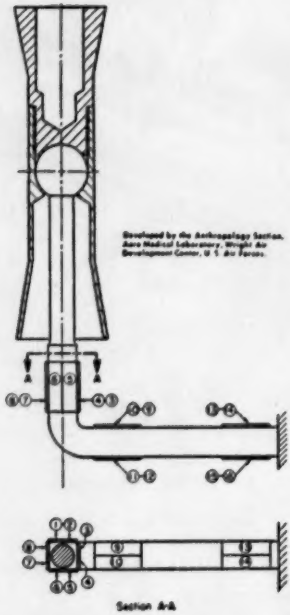

FIGURE X. Placement of Strain Gauges on Dynamometer.

Summary

In retrospect, the measuring of human strength has been the concern of scientists ever since the experiments of De La Hire over 250 years ago. The dynamometer, especially the spring steel type, has been the principal instrument utilized for strength testing. Other types of dynamometers include the mercurial and the pneumatic. The ergograph, the cable tensiometer, and the electrical strain gauge are other strength measuring instruments that have been developed.

Many various kinds of spring steel dynamometers have been developed since Régnier first introduced his crude device in 1807. Although slightly different in design, all of the spring steel dynamometers are based upon the same principle, namely deformation of a piece of steel either in the form of a ring, ellipse, or coil with the deformation of the metal being proportional to the force applied. In the United States, physical educators have generally used the same spring steel dynamometers developed by Sargent, psychologists the Smedley dynamometer, and orthopedists and physical therapists Martin's spring balance apparatus or its improved counterpart.

The Kellogg mercurial dynamometer is based upon the hydraulic principle. It has not been widely utilized, partly because of the relatively high cost for the machine and also becauses of its cumbersome size.

The pneumatic dynamometer is suitable only for measuring strength of grip. Its chief advantage lies in the fact that even the weakest grip will record a score.

The spring steel, mercurial, and pneumatic dynamometers together with the cable tensiometer were designed to measure a single maximum effort. In some instances, these instruments have been adapted so recordings could be made on a kymograph and the devices used in fatigue studies. The ergograph and the electrical strain gauge dynamometers were developed with the idea of securing a series of muscular efforts or for making a total record of the force from the initial movement through the drop-off. It should be pointed out that the cost of the electrical strain gauge type of instrument is higher than the spring steel dynamometer.

REFERENCES

1. Amar, Jules. *The Human Motor*, trans. Elsie P. Butterworth and George E. Wright. New York: E. P. Dutton and Co., 1920.
2. Bennett, Bruce L. *The Life of Dudley Allen Sargent, M.D., and His Contributions to Physical Education.* Unpublished doctoral dissertation, University of Michigan, 1947.
3. Binet, Alfred and Vaschide, N. Critique du dynomomètre ordinaire, *L'Année Psychologique*, **4**: 245-52, 1897.
4. Broca, P. *Instruction générales pour les recherches anthropologiques.* Paris: Victor Masson et fils, 1865.
5. Bullard, William N. Dynamometer, in Albert H. Buck (ed.) *Reference Handbook of the Medical Sciences*, Vol. II. New York: William Wood and Co., 1886.
6. Burq, D. Dynamometrie Médicale, *L'Union Médicale*, 2 ème série, **III**: 460-462 (Sept. 1859).
7. Carlson, L. D., and A. W. Martin, A Method for Studying Reflex Activity Under Varying Conditions, *Federation Proc.*, **7**: 18-19 (March 1948).
8. Carman, Ada. Pain and Strength Measurements of 1,507 School Children in Saginaw, Michigan, *American J. Psychology*, **10**: 392-98 (April 1899).
9. Clarke, H. Harrison, Objective Strength Tests of Affected Muscle Groups Involved in Orthopedic Disabilities, *Research Quarterly*, **19**: 118-147, 1948.
10. ———, Comparison of Instruments for Recording Muscle Strength, *Research Quarterly*, **25**: 398-411, 1954.
11. De La Hire, M. Examen de la force de l'homme, *Mémoires de l'Académie Royale des Sciences*, **12**: 153-62 (Nov. 14, 1699).
12. Dobie, W. B. and Peter C. G. Isaac. *Electric Resistance Strain Gauges.* London: English Universities Press Limited, 1948.
13. Galton, Francis. Range in height, weight, and strength of arm of males at each age," under Report of the Anthropometric Committee in *Reports of the British Association for the Advancement of Science 1881.*
14. Geckler, E. O. Pneumatic dynamometer, *American J. Surgery*, **45**: 353, 1939.
15. Hamilton, A. L. A new dynamometer, *The Psychological and Medico-Legal J.*, **2**: 255-56, 1875.
16. Hammond, G. M. A New Form of Dynamometer, *New York Medical J.*, **14**: 83, 1885.

17. HAMMOND, W. A. The dynamometer and dynamograph of Mathieu, *Quarterly J. of Psychological Medicine and Medical Jurisprudence*, **2**: 139-46, 1868.
18. HELLEBRANDT, F. A., *et al.* New devices for disability evaluation, *Archives of Physical Medicine*, **29**: 21-28, 1948.
19. HUNSICKER, PAUL. Arm strength at selected degrees of elbow flexion, *Air Force Project No. 7214-71727*, WADC TR 54-548 Wright Patterson Air Force Base, 1954.
20. KELLOGG, J. H. A new dynamometer for use in anthropometry, *Modern Medicine and Bacteriological World*, **2**: 269-75, 1893.
21. LAMBERT, E. H. Strain Gages: Resistance Wire, *Medical Physics*, Vol. 2, Edited by Otto Glasser. Chicago: Yearbook Publishers, 1950, pp. 1090-1098.
22. LOVETT, R. W., and E. G. MARTIN. The spring balance muscle test, *American J. Orthopedic Surgery*, **14**: 415-24, 1916.
23. MOSSO, ANGELO, Ueber die Gesetze der Ermudung Untersuchungen an Muskeln des Menschen, *Archiv für Anatomie und Physiologie* (1890), pp. 89-243.
24. POORE, G. V. *Selections from the Clinical Works of Dr. Duchenne (de Boulogne)*. London: The New Sydenham Society, 1883.
25. RALSTON, H. J., *et al.* Mechanics of human isolated voluntary muscle, *American J. Physiology*, **151**: 612-620, 1947.
26. SARGENT, D. A. Strength tests and the strong man of Harvard, *American Physical Education Review*, **2**: 108-119, 1897.
27. SCHMIER, A. A. Research work on a more precise method of determining muscle strength in poliomyelitis patients, *J. Bone and Joint Surgery*, **27**: 317-26, 1945.
28. SMEDLEY, FRED W. Report on child-study investigation, *Forty-Sixth Annual Report of the Chicago Board of Education for the Year Ending June 30, 1900.*
29. THOMSON, WILLIAM. On the electro-dynamic qualities of metals, *Philosophical Transactions of the Royal Society of London*, **146**: 649-751, 1856.
30. TUTTLE, W. Relation of maximum grip strength to grip strength endurance, *J Applied Physiology*, **2**: 663-670, 1950.
31. TUTTLE, W., *et al.* Relation of maximum back and leg strength to back and leg strength endurance, *Research Quarterly*, **26**: 96-106, 1955.
32. VERNON, H. M. *The Effects of Posture and Rest in Muscular Work*: Report *No. 29 of the Industrial Fatigue Research Board.* London: His Majesty's Stationery Office, 1924.

(Submited 11/1/54)

Aggression in Boxers and Wrestlers As Measured by Projective Techniques[1]

BURRIS F. HUSMAN
University of Maryland
College Park, Maryland

Abstract

In this study, the Rosenzweig P-F Study, selected TAT pictures, and a Sentence Completion Test were administered at intervals throughout the season to the following college groups: 9 boxers, 8 wrestlers, 9 cross country runners, and 17 control subjects. The tests were analyzed for number, severity, and direction of aggressive responses. Significant differences indicated that the boxers were least aggressive of the groups, and that they tended to direct their aggressive feelings inwardly (intra-punitive) rather than outwardly upon persons or things in their environment (extra-punitive). Indications were that the intensity and direction of aggression of these various athlete and non-athlete groups were quite different. The Thematic Apperception Test was judged the best instrument for assessing aggression.

IN THE PAST FEW decades there has been considerable research devoted to the validation of sports and other physical education activities as educational tools. Much of this research has been physiologically and kinesiologically oriented; little has dealt with the psychological aspects of physical education.

The Problem

The problem of this research was to study the aggression of college boxers and wrestlers before, after, and during the season; and before and after a contest. Specifically, the present research was designed:

1. To compare the aggression of boxers and wrestlers with each other, a control group (composed of cross country participants and non-athlete college students), and with a normal population.
2. To determine the effects of a season of competition and one emotionally charged match upon the aggression of the combative sport participants.
3. To shed further light upon the frustration-aggression theories proposed by Dollard, et al. (2), thus determining if sport participation lowers the aggression (Cathartic Theory) or if it increases the aggression (Circular or Inter-action Theory) (5).
4. To appraise the projective technique (a series of psychological tests, in which the subject is caught off guard, responds to an unstructured stimulus, and in so doing, divulges information about himself which he does not know or is not willing to admit) (1) as a tool for probing the personality traits of sport participants.

---

[1]This study was made in partial fulfillment of the requirements for a Doctor of Education degree in the Graduate College, University of Maryland, under the direction of Dr. Warren R. Johnson, College of Health, Physical Education, and Recreation.

The Procedure

The Rosenzweig Picture-Frustration Study (8), six selected pictures from Murray's Thematic Apperception Test (6), and a 20-item Sentence Completion Test (10, 11) were administered to 9 college boxers, 8 wrestlers, 9 cross country runners, and 17 control subjects under the following conditions.

1. Pre-season (T_1)—a week or more prior to the official opening of practice for each sport.
2. Pre-contest and mid-season (T_2)—about half-way through the season and less than two days prior to an intercollegiate match.[2]
3. Post-contest (T_3)—less than two days following an intercollegiate match.
4. Post-season (T_4)—one week or more following cessation of practice and the last intercollegiate match of the season.

The Rosenzweig P-F Study was scored for direction of aggression and type of reaction, as prescribed by Rosenzweig (9). Six of the 11 Rosenzweig scoring factors were analyzed in the results of this research. They were:

1. Extra-punitiveness—where the aggression is employed overtly and directed outwardly toward the environment.
2. Intra-punitiveness—where the aggression is turned by the subject upon himself.
3. Impunitiveness—where the aggression is evaded in an attempt to gloss over the frustration.
4. Ego-defense—where the ego of the subject predominates.
5. Extra-punitive ego-defense—where the subject aggresses extra-punitively to protect his ego.
6. Super-ego—where there is some accusation, charge or incrimination of the subject by someone else.

The Thematic Apperception Test (TAT) and the Sentence Completion Test (SCT) were scored for over-all intensity of aggression by counting and rating the severity of the aggressive words expressed (4) and for direction by determining if the aggression was turned toward some person other than the hero (extra-punitive aggression) or toward the hero (intra-punitive aggression).

Reliability of scoring was established for each of the three tests employed. An 85-per-cent agreement was obtained between the author and an independent scorer (who had previously worked with Rosenzweig) on the scoring of the Rosenzweig P-F Study. On the Thematic Apperception Test and the Sentence Completion Test, correlations were obtained between the author and an independent scorer on the scoring factors, these correlations being .90 for the factors of over-all intensity and extra-punitive aggression and between .60 and .77 for intra-punitive aggression.

Treatment of the data

The t test was utilized to determine if a significant difference existed between the means of the scoring factors (3). The means of the boxers were compared with those of the wrestlers, cross country runners, and control subjects on all four test situations (T_1, T_2, T_3, and T_4). Likewise, the means of

[2]The cross country participants were not tested prior to and after a contest.

the wrestlers were compared with those of the cross country runners, and control subjects; and the means of the cross country runners were compared with those of the control group.

Comparisons were also made between the means of each of the groups on the various test situations (T_1 T_2, T_3, and T_4), to study the effects of one contest and the season upon aggression.

The Results[3]

Significant differences between means at the 5-per-cent level of confidence or better indicated the following:

1. The boxers possessed less over-all intensity of agression than the wrestlers, cross country runners, and the control group as measured by the Thematic Apperception Test. The t's were 2.92, 3.40, and 2.89 respectively, all being significant at the 1-per-cent level of confidence or better with 25 degrees of freedom.

2. The Rosenzweig Picture-Frustration Study and the Sentence Completion Test (SCT) indicated that the boxers had less tendency than the cross country runners and the control subjects to express their aggression outwardly (extra-punitiveness), and that the boxers tended to blame themselves more for their frustrations or to gloss over their frustrations so that no person or object was to blame. Five t's, all significant beyond the 5-per-cent level of confidence, indicated this. The Thematic Apperception Test, however, indicated the boxers to possess less intra-punitive aggression than the wrestlers, cross country runners, and the control subjects. It was thought that the author's method of scoring intra-punitive aggression on this test may have been inadequate. This is partially verified by the low correlations obtained between the author and an independent worker on this scoring factor.

3. The cross country runners tended to aggress outwardly (extra-punitively) more than the boxers to protect their egos. The mean score of the boxers was 6.1 and that of the cross country participants 9.8. The t between these two means was 3.38; a t of 2.90 being significant at the 1-per-cent level of confidence with seventeen degrees of freedom.

4. On T_2 and T_4, the cross country runners possessed more ego defense than the boxers. The t's were 3.38 and 3.10 respectively; a t of 2.90 being significant at the 1-per-cent-level of confidence.

5. A t of 2.56 indicated that the boxers, after the contest (T_3), possessed more super-ego than the control group. A t of 2.06 with 25 degrees of freedom was required to establish confidence at the 5-per-cent level. It would be expected that the boxers would score higher on the trait super-ego after the

[3]Because of the large number of comparisons made, it is impossible to present the results of this study in brief tabular form. They have been presented in both tabular and graphical form in the author's dissertation, and are available for study by writing the author.

contest, since super-ego purports to assess the subject's guilt feelings for his overt action.

6. The cross country runners were more extra-punitive and less impunitive than the control group. A t of 2.98 on T_4 was significant at the 1-per-cent level of confidence, indicating the runners to be more extra-punitive; and a t of 2.50 on T_2 indicated that the control subjects were more impunitive. A t of 2.79 was significant at the 1-per-cent level of confidence and a t of 2.06 was significant at the 5-per-cent level, with 25 degrees of freedom.

7. By comparing Rosenzweig's trait of super-ego between the cross country group and the control group, a t of 2.53 was computed on T_4, which indicated that the runners possessed more super-ego than the control group. Again a t of 2.06 was needed to establish the 5-per-cent level of confidence. This significant difference between the cross country runners and the control group may be explained by noting that the cross country runners were more extra-punitive (see item 6 above)—thus aggressing against persons and objects in the environment more than the control group. Therefore, one would expect them to feel more guilty over their actions than the control subjects.

In addition to the significant differences described above, some non-significant differences seemed to suggest tendencies in and among the groups.

1. The boxers apparently possessed more of Rosenzweig's trait—super-ego—than the normal population[4] (7). The boxers' mean seemed to decrease after the season, and approached that of the normal population.

2. Indications were that an emotionally charged contest tended to lower the extra-punitiveness, ego-defense, and extra-punitive ego-defense means of the participants.

3. A combative contest tended to have the effect of causing the participants to turn their aggression inwardly.

4. A season of participation tended to increase the aggression of the participants as measured by the TAT.

5. A season of participation caused the Rosenzweig scoring factors of extra-punitiveness, intra-punitiveness, and super-ego to decrease for the boxers and cross country runners.

6. The results of this research tended to substantiate both the cathartic and circular theories of aggression.

7. Of the 23 significant differences found in this research, 13 were found by analyzing the TAT data. On this basis, the TAT seemed to be the best instrument for assessing aggression.

8. The TAT may be scored objectively by counting the frequency of aggressive words and rating their intensity.

---

[4]The normal population here refers to the population used by Dr. Saul Rosenzweig in norming the Rosenzweig P-F Study.

Conclusions and Recommendations

The results of this study indicated that the aggression of boxers was significantly different from that of the wrestlers, cross country runners, and control subjects. The boxers possessed less over-all intensity of aggression, were less extra-punitive and more intra-punitive and impunitive than the other subjects. The cross country runners were more extra-punitive and less intra-punitive and impunitive than the control subjects.

Trends in the data which were not significant indicated that a season of participation tended to increase the aggression of the participants as measured by the Thematic Apperception Test, while a season of participation tended to lower the Rosenzweig scoring factors of extra-punitiveness, intra-punitiveness, and super-ego. An emotionally charged contest tended to lower the Rosenzweig scoring factors of extra-punitiveness, ego-defense, and extra-punitive ego-defense, and seemed to cause the participants to turn their aggression inwardly.

This research tended to substantiate both the cathartic and circular theories of aggression and indicated that the projective technique is useful in this type of research. Future studies might well be conducted using the TAT or Rosenzweig P-F to study aggression throughout a season or before and after a contest. The sport situation presents an unusual opportunity for analysis of personality traits. In sports, real life situations exist which offer physical educators and psychologists an ideal environment in which to conduct research.

REFERENCES

1. Clark, R., A Method of Administering and Evaluating the Thematic Apperception Test in Group Situations, *Genetic Psychology Monogram*, **30**: 3-55, Aug. 1944.
2. Dollard, J., N .E. Miller, L. W. Doob, O. H. Mowrer, and R. R. Sears et al., *Frustration and Aggression.* New Haven: Yale University Press, 1939. 150 pp.
3. Guilford, J. P., *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill Book Co., Inc., 1950. 633 pp.
4. Lindzey, Gardner, An Experimental Test of the Validity of the Rosenzweig Picture Frustration Study, *Journal of Personality*, **18**: 315-320, Mar. 1950.
5. Morlan, George K., A Note on the Frustration-Aggression Theories of Dollard and his Associates, *The Psychological Review*, **56**: 1-8, Jan. 1949.
6. Murray, Henry A., M.D., and Staff of Harvard Psychological Clinic, *Thematic Apperception Test Manual.* Cambridge, Mass.: Harvard University Printing Office, 1943. 20 pp.
7. Rosenzweig, Saul, Revised Norms for the Adult Form of the Rosenzweig Picture-Frustration Study, *Journal of Personality*, **18**: 344-346, Mar. 1950.
8. ———, The Picture Association Method and Its Application in a Study of Reactions to Frustrations, *Journal of Personality*, **14**: 3-23, Sept. 1945.
9. ———, Edith Fleming, and Helen J. Clark, Revised Scoring Manual for the Rosenzweig Picture-Frustration Study, *Journal of Psychology*, **24**: 165-208, Oct. 1947.
10. Stein, Morris, The Use of a Sentence Completion Test for the Diagnosis of Personality, *Journal of Clinical Psychology*, **3**: 47-56, Jan. 1947.
11. Symonds, P. M., The Sentence Completion Test as a Projective Technique, *Journal of Abnormal and Social Psychology*, **42**: 320-329, July 1947.

(Submitted 1/8/55)

Relationship Between Strength Loss in Pack Carrying and Certain Motor-Physical Fitness Criteria[1]

DONALD K. MATHEWS
Washington State College, Pullman, Washington
CLAYTON T. SHAY
Springfield College, Springfield, Massachusetts
and
H. HARRISON CLARKE
University of Oregon, Eugene, Oregon

Abstract

The purpose of this study was to determine the relationship between strength loss from carrying military packs under field conditions and various motor-physical fitness tests. The subjects were conditioned from six marches carrying the combat load of the rifleman. On the final march, the subjects carried 61 pounds with rucksack for 7.5 miles at a rate of 2.5 miles in 50 minutes. The highest multiple correlation obtained was —.92 between strength loss and the Strength Index; similar correlation with the Army, Navy, and AAF motor fitness tests were between —.57 and —.49. Individual test items of special significance in this study were leg-lift strength, time in the 300-yard run, and number of squat thrusts in one minute. For most effective pack carrying, individuals with high Strength Indices should be utilized.

DURING WORLD WAR II, various branches of the Armed Forces found it necessary to devise tests which would measure the fitness of personnel to perform physical assignments. Tests were proposed which could be easily administered in terms of time, equipment, and unskilled examiners. These tests were extensively used in physical training programs throughout the war.

In constructing these tests, validity was based largely upon hypothetical reason rather than upon criteria involving the arduous duties Army, Navy, and Air Force personnel must accomplish in line of duty. In all instances, the basic procedures in validating these tests were as follows: (1) empirical decision to include such motor and physical aspects of fitness as strength, endurance, speed, agility, muscular explosiveness, and the like; (2) theoretical selection of test items, such as chins, dips, runs, etc., to test the motor fitness aspects listed; and (3) relation of the individual test items, singly and in combination, to a composite of the larger number of test items. Thus, for

[1]This study was conducted at Springfield College under subsidy from the Climatic Research Laboratory, Quartermaster Corps, Department of the Army, Natick, Mass. Acknowledgments are expressed to the following graduate assistants who helped in the collection and analysis of the necessary test data: Gavin H. Carter, David H. Clarke, Karl K. Klein, William Mahon, Barry Thompson, and James E. Torpey.

example, with the Army "Physical Efficiency Test," 23 test items were included in the initial experimentation; today, the test battery is composed of five of these items. Evidence as to whether these test batteries measure the type of physical condition required of the rifleman (infantry soldier) is inconclusive.

The purpose of this study is to determine if the armed services motor fitness tests are correlated to one realistic military requirement, the soldier's endurance in carrying military packs. In addition, certain other motor-physical fitness tests in use in civilian physical education programs were included to determine their value for use in this military situation. These relationships were studied when the subjects were not trained for long marches; and again following a series of six marches, one week apart, the subjects thus being conditioned for marching.

Research Procedures

1. *Motor-Physical Fitness Tests.* The motor-physical fitness tests included in this study, with brief descriptions of each,[2] were as follows:

Army Physical Efficiency Test (13). This test is designed to measure the principal factors in general motor fitness designated as strength, endurance, agility, and co-ordination. The present test items are: pull-ups, squat-jumps, push-ups, unlimited sit-ups, and 300-yd. shuttle run.

Navy Standard Physical Fitness Test (11). The Navy test consists of the following five items: squat thrusts, sit-ups, push-ups, squat-jumps, and pull-ups.

Air Forces Physical Fitness Test (1). Three test items compose the AAF motor fitness battery: sit-ups, pull-ups, and 300-yd. shuttle run.

Strength Index (12). The Strength Index (SI) is the gross score obtained from lung capacity and the following six strength tests: right grip, left grip, back lift, leg lift, pull-ups, and push-ups (the latter two tests combined in a formula provide an arm strength score). The SI is considered in physical education as a measure of general athletic ability.

Physical Fitness Index. The Physical Fitness Index (PFI) is a score derived from comparing an achieved Strength Index with a norm based upon the individual's sex, weight, and age. This test is proposed as a measure of general physical fitness, indicating the immediate capacity of the individual for physical activity.

McCloy Arm Strength Score (10). The formula for the McCloy Arm Strength Score is as follows:

$$\text{Chinning or Dipping Strength} = 1.77\,(\text{Weight}) + 3.42\,(\text{No. of Chins or Dips}) - 46.$$

In this study, push-ups and pull-ups were combined in the formula. McCloy demonstrated the importance of arm strength in athletic ability.

---

[2]Details concerning all motor-physical fitness test batteries are contained in reference 6.

Over-All Strength Loss. This criterion measure was obtained by adding the strength losses of all muscle groups for a given march. In a sense, then, this score represents the over-all fatigue effect on muscular strength from the pack-carrying marches.

2. *Test Administration.* All strength and motor-physical fitness tests were administered before and after the series of pack-carrying marches. The testing techniques were in accordance with the standard procedures described in the references.

3. *Strength Decrement Index.* The fatigue effect of the military marches was determined from the strength losses of the primary muscle groups involved in the marches. In another study by the authors (8), the hypothesis was verified that muscle fatigue is accompanied by loss in the ability of the muscles to apply tension; the amount of fatigue can be identified as the proportionate loss in strength after fatiguing exercise, designated as the Strength Decrement Index (SDI). The SDI is computed as follows:

$$SDI = \frac{S_b - S_a}{S_b} \times 100$$

in which: S_b = Before march strength
S_a = After march strength

4. *Selection of Muscle Groups.* The muscle groups tested for strength loss were those primarily involved in supporting the pack, in stabilizing the body in relation to the pack, and in marching.[3] Eight muscle groups[4] were included, as follows: shoulder elevators, trunk flexors, trunk extensors, hip flexors, hip extensors, knee flexors, knee extensors, and ankle plantar flexors. Clarke's cable-tension tests (5) were utilized to measure the strength of these muscle groups.

5. *Field Course.* The marching distance was 7.5 miles over moderately rough terrain. The course is 1.25 miles long, so was traversed six times for each march. The heavy broken line indicates the trail used for the field course.

6. *Rate of March.* The basic unit of time selected was patterned after the military concept of 50 minutes' marching followed by ten minutes of rest. The rate of march chosen was 2.5 miles, twice around the course, in 50 minutes. This rate, 88 yards per minute, conforms well with metabolic studies which have reported that the maximum efficiency of energy cost in marching is between 80 and 95 yards per minute (3, 4).

---

[3]For assistance in the selection of the muscle groups, acknowledgments are made to Dr. F. A. Hellebrandt, Head, Department of Physical Medicine and Rehabilitation, University of Illinois Research and Education Hospitals, and to Dr. Earl C. Elkins, Section on Physical Medicine, Mayo Clinic.

[4]Other muscle groups were actually included in the study, but did not show significant losses in strength from the marches, so are eliminated from this report.

An electric timing device was constructed which sounded four horns for three seconds every five minutes. Markers were placed on the trail at points the subjects should reach in each successive period of five minutes. In marching, if the subject was short of the proper marker when the horns sounded, he increased the pace a bit; if he had passed the marker, he slowed appropriately. This method of regulating the rate of march proved entirely satisfactory, as the subjects readily learned to pace themselves to reach the markers within 15 seconds before or after the sounding of the horns.

7. *Clothing and Equipment.* The subjects were provided with regulation Army clothing, combat boots, equipment, packs, and M-1 rifles. They were dressed in accordance with Army regulations. The load was distributed on the person in the manner it would be carried in the field. The packs, however, were varied as to kind, amount of weight, and location on the body, as will be described below.

8. *The Marches.* There was a total of seven marches in which the subjects participated during the total pack-carrying experiment (7). For purposes of this study, however, two marches were selected: the third march (the first when the subjects carried complete combat packs), when the subjects were relatively unconditioned for marching; and the seventh march, when the subjects were in their best marching condition. The specifications for the two marches were as follows:

Third March: Combat pack, weight, 41 lb.; pack carried high on back; winter clothing.
Seventh March: Rucksack; weight, 61 lb.; summer fatigues.

9. *Subjects.* The subjects were 32 male students at Springfield College chosen from the student population. Inasmuch as the sampling was composed predominantly of physical education majors, it was felt that their performance in the study would be superior where sustained physical effort was required. As a consequence, the following median scores further describe this group: height, 70 in.; weight, 168 lb.; Physical Fitness Index, 104; and Strength Index, 3013. The age range was between 18.5 and 26.5 years; the median age was 22.2.

10. *Testers.* Prior to the actual collection of test data, the graduate students serving as research assistants were carefully trained in testing techniques. They were required to demonstrate testing competence by obtaining objectivity coefficients by the test-retest method with different testers approximately equal to those established for the tests. The correlations thus achieved ranged from .86 to .98.

11. *Conditioning Effects.* The conditioning effects resulting from the seven pack-carrying marches were determined by comparing the means on all cable-tension strength and physical-motor fitness tests before the first march and after the last march. These differences were tested for significance using the t ratio. For 32 cases, t's of 1.70 and 2.04 are significant at the .05 and .01 levels of confidence respectively (9).

12. *Relationships.* For each march, the Strength Decrement Index for those strength tests exhibiting a significant loss from pack carrying was computed for the subjects. These SDI's were then correlated with each other and with each of the motor fitness tests by means of the Pearson product-moment method. In interpreting these correlations, r's of .35 and .45 are significant at the .05 and .01 levels of confidence respectively (9). The correlations of the strength losses with the test batteries are largely negative, as would be expected from this situation. In explanation, the SDI is reported as a positive figure, but actually indicates the proportionate amount of strength loss; thus, the desired result would be a low SDI related to a high test score.

Multiple correlations were computed by the Wherry-Doolittle method with each of the motor-physical fitness tests serving as criterion in turn. Inasmuch as statistical significance for multiple correlations is dependent in part upon the number of variables involved, the following listings are presented for purposes of interpretation.

Number of Variables	Degrees of Freedom	Coefficient of Multiple Correlation	
		.01 level	.05 level
3	29	.43	.51
4	28	.48	.56
5	27	.51	.59
6	26	.55	.62

Predictive indices (2) were utilized to indicate the percentage better than chance relationships for the various correlations obtained.[5] This statistic is a convenient means of roughly comparing the predictive significance of different coefficients of correlation, especially when N is constant, as in this study.

13. *Other Relationships.* In order to better understand the various motor-physical fitness tests, the interrelationships of the motor-physical fitness test batteries and items were determined for the last march, when the subjects were in their best condition.

Some caution should be exercised in interpreting the statistical significance of the results of this study, inasmuch as the subjects were somewhat superior to the population. Hence, they did not conform strictly to the normality of a random sample upon which reliability formulas are based.

Conditioning Effects from Marching

In an earlier report by the writers (7), the conditioning effects resulting from the seven pack-carrying marches were presented. In summary, the results were as follows:

1. The increases in mean strength of the involved muscles were definite: three of these were significant beyond the .01 level of confidence (trunk extensors, trunk flexors, and shoulder elevators); the gain in strength for the

[5]The Predictive Index is the reverse of Truman Kelly's coefficient of alienation. The formula: $1 - \sqrt{1 - r^2}$

hip flexors was nearly significant at this level; and the increase for ankle plantar flexors was noticeable, but did not reach the .05 level.

2. The means for all motor-physical fitness tests increased during the pack-carrying marches. All three of the armed forces motor fitness batteries showed definite gains, especially the Army Physical Efficiency Test, which had a t of 6.1. The mean Physical Fitness Index rose from 104 to 113. The mean increase in Strength Index was nearly significant at the .01 level; and the gain for McCloy's Arm Strength Score nearly reached the .05 level.

Relationships: Non-Conditioned Subjects (Third March)

PRODUCT-MOMENT CORRELATIONS

The product-moment intercorrelations between the experimental variables (SDI's) and the correlations of these variables with motor-physical fitness batteries as criteria for the relatively non-conditioned subjects on the third march appear in Table 1. Appropriate interpretation of these results follows.

TABLE 1

Intercorrelations of Strength Decrement Indices and Their Correlations with the Various Motor-Physical Fitness Criteria: Non-Conditioned Subjects

Item	Trunk Flex.	Hip Flex.	Hip Ext.	Knee Flex.	Knee Ext.	Ankle Plantar Flex.
Experimental Variables						
Trunk Flex.		−.04	−.40	−.02	−.29	−.09
Hip Flex.			−.08	−.22	−.34	.02
Hip Ext.				.03	.23	.00
Knee Flex.					.19	.00
Knee Ext.						.30
Motor-Physical Fitness Criteria						
Army Test	−.02	.15	−.14	−.59	.03	−.03
Navy Test	.00	−.08	−.27	.02	.24	.22
Air Force Test	−.05	.16	−.09	−.47	−.04	.01
Strength Index	.00	−.02	−.28	−.17	.05	−.08
Physical Fitness Index	−.08	.16	−.14	−.30	.01	−.14
McCloy Arm Strength	−.02	.19	−.17	−.19	.53	−.09
Composite SDI	.09	.26	.60	.35	.53	.26

Intercorrelations of SDI's. The intercorrelations of the SDI's for this march were low. The highest such correlation was −.40 between trunk flexion and hip extension SDI's, which is statistically significant between the .05 and .01 levels of confidence. A second correlation, −.34 between hip flexion and knee extension SDI's, is significant at the .05 level.

Correlations with motor-physical fitness tests. Four of the correlations with motor-physical fitness criteria were significant at and beyond the .01 level. Two of these correlations were with knee flexor SDI: −.59 with the Army test, and −.47 with the AAF test. The other two correlations significant at

this level were obtained with composite SDI: .60 with hip extension SDI, and .53 with knee extension SDI. Also, knee flexion SDI was significantly correlated with the composite SDI at the .05 level (.35).

MULTIPLE CORRELATIONS

With the exception of the composite SDI criterion, none of the multiple correlations obtained for this march exceeded the highest product-moment correlations for the various motor-physical fitness batteries. A multiple correlation of .95 resulted from the SDI's for the muscle groups and the composite SDI. This multiple is spuriously high, since the SDI elements were involved in both factors correlated. The SDI's were for the following muscle groups: hip extensors, ankle plantar flexors, trunk flexors, and knee extensors.

Relationships: Conditioned Subjects (Seventh March)

PRODUCT-MOMENT CORRELATIONS

For the conditioned subjects participating in the final march of the study, the intercorrelations between the SDI's for the various muscle groups and the correlations of these variables with the motor-physical fitness criteria appear in Table 2.

TABLE 2

Intercorrelations of Strength Decrement Indices and Their Correlations with the Various Motor-Physical Fitness Criteria: Conditioned Subjects

Item	7 Shldr. Elev.	1 Trunk Flex.	3 Trunk Ext.	2 Hip Flex.	4 Hip Ext.	6 Knee Flex.	5 Ankle Plantar Flex.
Experimental Variables							
7 Shldr. Elev.		.12	.09	.00	−.20	.18	.23
1 Trunk Flex.			.09	.10	.18	−.18	.03
3 Trunk Ext.				−.07	.10	.42	.09
2 Hip Flex.					−.11	.10	.28
4 Hip Ext.						−.09	.09
6 Knee Flex.							.16
Motor-Physical Fitness Criteria							
Army Test	.03	.25	−.03	−.20	.31	−.43	.17
Navy Test	.14	.23	−.16	−.12	.33	−.45	.12
Air Force Test	−.04	.17	−.09	−.19	.24	−.49	.07
Strength Index	−.06	.01	.03	−.20	−.59	−.53	−.16
Physical Fitness Index	.12	.12	.02	−.24	−.03	−.40	−.03
McCloy Arm Strength	−.39	−.15	.00	−.10	−.24	.20	−.14
Composite SDI	.40	.44	.36	.44	.24	.47	.50

Intercorrelations of SDI's. For the conditioned subjects, only one of the intercorrelations between the SDI's was significant, the .42 for trunk extension and knee flexion, which nearly reaches significance at the .01 level of confidence.

Correlations with motor-physical fitness tests. Nine of the correlations between SDI's and motor-physical fitness criteria were significant near the .01

level of confidence and beyond. Five correlations were significant, or nearly so, at the .05 level. The highest correlation was —.59 between hip extension SDI and the Strength Index.

The most significant feature of this correlational matrix is the prominence of knee flexion SDI. For all but one criterion, the McCloy Arm Strength Score, the correlations with this variable were significant beyond the .05 level (ranging from —.40 to —.53). There is a definite pattern indicating less strength loss of the knee flexor muscles from pack carrying when the subjects are in good physical condition, as determined by several motor physical-fitness tests, and when the subjects themselves have been conditioned to marching.

Five other significant correlations were obtained with composite SDI, as follows: .50 with ankle plantar flexion SDI; .44 with both trunk flexion and hip flexion SDI's; .40 with shoulder elevation SDI; and .36 with trunk extension SDI. A correlation of —.39 was also obtained between the McCloy Arm Strength Score and shoulder elevation SDI.

MULTIPLE CORRELATIONS

The distribution of experimental variables (SDI's) as they appeared in the various multiple correlations with motor-physical fitness criteria for the final march is consolidated in Table 3.

TABLE 3

*Consolidation of Multiple Correlations with Motor-Physical Fitness Tests for the Final March: Conditioned Subjects**

SDI's	SI	PFI	Army Test	Navy Test	AAF Test	McCloy Arm Strength	Composite SDI's
Shoulder Elevation				3		1	6
Trunk Flexion	1						2
Trunk Extension	3						
Hip Flexion	4	2					4
Hip Extension			2	2	2	2	5
Knee Flexion	2	1	1	1	1	3	3
Ankle Plantar Flexion							1
Multiple R	—.92	—.42	—.49	—.57	—.51	—.49	.88
Predictive Index	.61	.10	.13	.19	.14	.13	.53

*Numbers on chart refer to order in which SDI's appear in the multiple correlations.

Conclusions which may be drawn from this table are as follows:

1. The highest multiple correlation was —.92 with the Strength Index as the criterion. This very high inverse relationship indicates definitely that the more gross strength an individual possesses the less proportionate strength loss he will experience from pack-carrying marches up to 7.5 miles.

2. The multiple correlation with Physical Fitness Index is much lower (—.42); the predictive value for SI is six times as great as for PFI. Thus, the relationship between proportionate body strength (for age and weight) and strength loss is low, being statistically significant at the .05 level of confidence.

3. The highest multiple correlation for the Armed Forces test was —.57, when the Navy test was the criterion; the AAF and Army tests were nearly the same (—.51 and —.49 respectively). Contrasting predictive indices, the SI had 3.2 times greater predictive value than did the Navy test, while the Navy test exceeded the PFI by 90 per cent.

4. The multiple correlation utilizing the McCloy Arm Strength Score as criterion was —.49, the same as for the Army test.

5. The multiple R with composite SDI as the criterion measure was high (.88). This occurrence was expected since the individual SDI's, themselves, composed the criterion.

6. The following SDI's were of greatest importance in the various multiple batteries:

(a) Knee flexion ranked first in four multiples; second, in one; and third, in two.
(b) Hip extension appeared second in four multiples; and fifth, in one.
(c) Other SDI's ranking first in the multiples were: trunk flexion, for Strength Index; shoulder elevation, for McCloy Arm Strength Score; and ankle plantar flexion, for composite SDI.

7. Because of the high multiple correlation of —.92 when the Strength Index was the criterion, a multiple regression equation was computed, as follows:

$$X_e = -.47X_1 - .63X_2 + .40X_3 + .02X_4 + 3402$$

where: X_e = Strength Index
X_1 = Hip Extension SDI
X_2 = Knee Flexion SDI
X_3 = Trunk Extension SDI
X_4 = Hip Flexion SDI

Intercorrelations of Motor-Physical Fitness Tests

TEST BATTERIES

The test scores obtained on each of the motor-physical fitness test batteries after the last march, when the subjects were in a conditioned state, were intercorrelated. Table 4 contains the results of this analysis.

Relationships among the service tests revealed high intercorrelations of the Army test with the other two: .88 with the Navy test; and .86 with the AAF battery. The correlation between the AAF and the Navy tests, however, was only .48. These tests have common test elements, as follows: pull-ups and sit-ups appear in all three batteries; push-ups and squat jumps, in the Army and Navy tests; 300-yd. shuttle run, in the Army and AAF tests; and squat-thrusts, in the Navy test only. The Army test has the highest communality with the other two batteries, which may account for its high relationship with each of them. The Army test also correlated .80 with the PFI.

TABLE 4

Intercorrelation of Motor-Physical Fitness Test: Conditioned Subjects

Test	Rogers Arm Str.	SI	PFI	Navy Test	Army Phys. Effic. Test	AAF Test
McCloy Arm Strength	.46	.66	.14	.10	—.04	.14
Rogers Arm Strength		.86	.80	.34	.59	.69
Strength Index			.83	.32	.36	.50
Physical Fitness Index				.49	.80	.63
Navy Test					.88	.48
Army Phys. Effic. Test						.86
AAF Test						

The Strength Index correlated .86 with the Rogers Arm Strength Score; and .83 with the Physical Fitness Index. This arm strength score is a part of the SI; and the PFI is obtained from the SI.

The Rogers Arm Strength Score resulted in considerably higher correlations with the various test batteries than did the McCloy Arm Strength Score. The Rogers score correlations with the Navy, Army, and AAF tests were .34, .59, and .69 respectively; while for the McCloy score, these correlations were .10, —.04, and .14. The correlation between the two arm scores was .46.

TEST ITEMS

Table 5 contains the intercorrelations of six of the test items composing the physical-motor fitness tests and the correlations of these items with the various batteries. The spuriousness of certain of these correlations is admitted, when the test item is also a part of a test battery, as in leg lift strength vs. Strength Index and number of sit-ups vs. Army, Navy, or AAF test.

TABLE 5

Intercorrelations of Motor-Physical Fitness Items and Their Correlations with the Various Motor-Physical Fitness Criteria: Conditioned Subjects

Item	Leg Lift	Back Lift	Sit-Ups	300-Yd. Shuttle	Squat Thrusts	Squat Jumps
Experimental Variables						
Leg Lift		.549	—.011	—.570	.029	.072
Back Lift			—.023	—.445	—.109	—.280
Sit-ups				—.261	.482	.483
300-Yd. Run					—.409	—.149
Squat Thrusts						.637
Squat Jumps						
Motor-Physical Fitness Criteria						
AAF Test	.416	.342	.554	—.742	.507	.268
Navy Fitness	.264	.095	.606	—.605	.793	.756
Army Phys. Effic. Test	.341	.129	.632	—.811	.623	.540
SI	.936	.597	.037	—.718	.343	—.093
PFI	.816	.458	—.160	—.658	.437	.074
McCloy Arm Strength	.507	.642	—.149	—.457	—.114	—.202
Rogers Arm Strength	.656	.493	—.046	—.731	.306	.090

The highest intercorrelatons were with leg lift strength: —.57 with time in the 300-yd. shuttle run; and .55 with back lift strength. Let lift strength also correlated high with the following motor-physical fitness criteria: .94 with SI; .82 with PFI; and .66 with Rogers Arm Strength Score.

Time in the 300-yd. shuttle run had significant correlations with all but one of the variables included in Table 5. These correlations were —.81 with Army test; —.74 with AAF test; —.73 with Rogers Arm Strength Score; —.72 with SI; —.66 with PFI; —.61 with Navy test; —.57 with leg lift; —.46 with McCloy Arm Strength Score; —.45 with back lift strength; and —.41 with squat thrusts.

The number of squat thrusts performed in one minute was significantly related to the following variables: .79 with Navy test; .64 with squat jumps; .62 with Army test; .51 with AAF test; .48 with unlimited sit-ups; .44 with PFI; —.41 with 300-yd. shuttle run; and .34 with SI.

Significant correlations with back lift strength were: .64 with McCloy Arm Strength Score; .60 with SI; .55 with leg lift strength; .49 with Rogers Arm Strength Score; —.45 with 300-yd. shuttle run; and .46 with PFI.

For sit-ups, the significant correlations were: .63 with Army test; .61 with Navy test; .55 with AAF test; and .48 with both squat thrusts and squat jumps.

Four correlations were significant for squat-jumps: .76 with Navy test; .64 with squat thrusts; .54 with Army test; and .48 with sit-ups.

MULTIPLE CORRELATIONS

Multiple correlations were computed utilizing the motor-physical fitness tests as criterion measures and six of the items comprising these tests batteries as independent variables. A summary of these correlations appears in Table 6.

TABLE 6

Consolidation of Multiple Correlations with Motor-Physical Fitness Tests and Motor-Physical Fitness Items for Final March: Conditioned Subjects

Item	AAF	Navy	Army	SI	PFI
Leg Lift				1	1
Back Lift					
Sit-ups	2		2		3
300-Yd. Run	1	3	1		
Squat-Thrusts		1		2	2
Squat-Jumps		2	3		
Multiple R	.82	.93	.95	.98	.99

As would be expected, the resultant multiple correlations were very high. The two highest correlations were .99 and .98 for the PFI and SI respectively. In both of these multiples leg lift strength and squat thrusts were the number 1 and 2 variables; sit-ups was included in the PFI sequence.

For the service tests, the correlations were .95, .93, and .82 for the Army, Navy, and AAF tests in order. The number 1 variable for these multiples was 300-yd. shuttle run for the Army and AAF tests, and squat thrusts for the Navy test; the number 2 variable was sit-ups for the Army and AAF tests, and squat-jumps for the Navy test. A third variable was included in two of the multiples: squat jumps for the Army test; and 300-yd. shuttle run for the Navy test.

Summary and Conclusions

The purpose of this study was to determine the relationship between strength loss from carrying packs under field conditions and various motor-physical fitness tests. Thus, these tests were validated against one strenuous duty assignment of the rifleman, marching distances with full combat load. It should be clearly understood that no conclusions are possible from this research relative to the validity of these tests for other physically strenuous military assignments.

While in this report emphasis has been placed on the contrast between the effectiveness of conditioned and non-conditioned subjects in carrying packs, two other factors should be kept in mind in comparing the results of the two marches, as follows: (a) the load carried by the conditioned group was 20 pounds heavier than for the non-conditioned subjects; and (b) a rucksack was carried by the conditioned subjects; the non-conditioned subjects carried a regulation combat pack, load in high position. A summary of the results follows:

1. Physical conditioning of the subjects definitely took place during the pack-carrying experiment. Evidences of this conclusion are as follows: (a) the strength of the muscle groups tested increased as the study progressed; (b) the overall percentage of loss in strength lessened and occurred later on the final march than on the earlier march, even though the final march was more strenuous; and (c) statistically significant increases were found in all but one of the motor-physical fitness tests at the termination of the research.

2. The intercorrelations of SDI's for both marches were generally low, none of them quite reaching significance at the .01 level. This result indicates that strength loss in carrying packs on marches with different packs and loads is specific to muscle groups, rather than consistent to all muscle groups primarily involved.

3. For the conditioned subjects on the final march, nine of the product-moment correlations of SDI's with motor-physical fitness tests were significant at the .01 level and beyond; five additional correlations were significant at the .05 level. For the non-conditioned subjects on the initial march, four such correlations were significant at the .01 level, with one other at the .05 level. Within the limited magnitude of these correlations, these relationships indicate that a high fitness score is accompanied by less strength loss and vice versa.

4. With the exception of the composite SDI, multiple correlations between the SDI's and the different motor-fitness tests were generally very low for the non-conditioned subjects. As a consequence, only meager conclusions are possible in this situation. Two things may account for these results: (a) the load and conditions may not have placed the subjects under sufficient stress for an adequate evaluation of the effects; and (b) the variations in the marching condition of the subjects were such that the effects may have been erratic from subject to subject, thus resulting in lowered correlations.

The final march placed the subjects under greater stress as they carried 61 pounds, an increase of 50 per cent in weight, and they were definitely improved in physical condition. As a consequence of these considerations, greater importance can be placed on the results of the final march in portraying the relationship between strength loss in carrying Army packs and the motor-physical fitness tests. Comments below relate to the conditioned subjects only.

5. The Strength Index as criterion produced by far the highest multiple correlation (—.92) obtained for any of the motor-physical fitness tests (excluding composite SDI). Consequently, the Strength Index is definitely superior to the other tests investigated in predicting muscle fatigue, as reflected in strength loss, from carrying military packs under the field conditions of this experiment.

6. Of the Armed Forces motor fitness batteries, the Navy test was found superior to either the Army or the AAF tests in evaluating strength loss from pack carrying. However, the multiple correlation obtained for this test was far too low for predictive purposes (—.57).

7. The McCloy Arm Strength Score produced a multiple correlation within the same range as the armed services tests. The Physical Fitness Index had a lower multiple than McCloy's score.

8. The individual SDI's, which were most prominent in the multiple correlations, were knee flexors and hip extensors. Apparently, therefore, the strength loss of these muscle groups from pack-carrying marches best reveals the general condition of the subjects. However, the muscle groups reflecting overall strength loss are ankle plantar flexors, trunk flexors, and knee flexors, as these had ranking positions in the multiple correlation with the composite SDI.

9. The Army motor fitness test correlated well with the other two service tests for conditioned subjects, as follows: .88 with the Navy test; and .86 with the AAF battery. The correlation between the Navy and AAF tests, however, was much lower (.48). The Strength Index correlated .86 with Rogers Arm Strength Score; and .83 with the Physical Fitness Index. The Rogers Arm Strength Score resulted in considerably higher correlations with the various test batteries than did the McCloy Arm Strength Score.

10. The three most significant tests resulting from an intercorrelation of six of the individual tests composing the motor-physical fitness batteries and

the correlation of these items with the various batteries were as follows:

(a) Leg lift strength: .94 with Strength Index; .82 with Physical Fitness Index; and .66 with Rogers Arm Strength Score.

(b) Time in the 300-yd. run: —.81 with Army test; —.74 with AAF test; —.73 with Rogers Arm Strength Score; —.72 with Strength Index; —.66 with Physical Fitness Index; and —.61 with Navy test.

(c) Number of squat thrusts in one minute: .79 with Navy test; .65 with squat jumps; and .62 with Army test.

11. With subjects conditioned for marching with packs, the following multiple correlations were obtained between each of the motor-physical fitness tests as criterion and six of the items comprising these batteries:

(a) Physical Fitness Index, .99: leg lift, squat thrusts, and sit-ups.

(b) Strength Index, .98: leg lift and squat thrusts.

(c) Army Physical Efficiency Test, .95: 300-yd. shuttle run, sit-ups, and squat jumps. (.92 without squat jumps.)

(d) Navy Standard Physical Fitness Test, .93: squat-thrusts, squat-jumps, and 300-yd. shuttle run (.85 without run).

(e) Air Forces Physical Fitness Test, .82: 300-yd. shuttle run and sit-ups.

12. For most effective pack carrying, individuals with high Strength Indices should be utilized.

REFERENCES

1. AAF Regulation No. 50-10, April 1943, Physical Fitness Test.
2. Bailer, E. M., *Content and Form in Tests of Intelligence.* New York: Teachers College, Columbia University, 1925, pp. 25-29.
3. Breznia, E., and H. Reichel, Energy Exchange During Marching. (1) Marching upon a Horizontal Road, *Biochem. Zeitschr.*, **63**: 170 (1914).
4. Cathcart, E. P., D. T. Richardson, and W. C. Campbell, The Maximum Load To Be Carried by the Soldier, *Journal of the Royal Army Medical Corps*, **40**: 435; 41-87 (1923).
5. Clarke, H. Harrison, *A Manual: Cable-Tension Strength Tests.* Chicopee, Mass.: Brown-Murphy Co., 1953.
6. ———, *Application of Measurement to Health and Physical Education*, 2nd Ed. New York: Prentice-Hall, Inc., 1950.
7. ———, Clayton T. Shay, and Donald K. Mathews, Strength Decrement from Carrying Various Army Packs on Military Marches of 7.5 Miles, *Research Quarterly*, **26**: 253-265 (Oct. 1955).
8. ———, Clayton T. Shay, and Donald K. Mathews, Strength Decrement of the Elbow Flexor Muscles Following Exhaustive Exercise, *Archives of Physical Medicine and Rehabilitation.* **35**: 560 (Sept. 1954).
9. Garrett, Henry E., Statistics in Education and Psychology, 4th ed. New York: Longmans, Green and Co., 1953.
10. McCloy, C. H., A New Method of Scoring Chinning and Dipping, *Research Quarterly*, **2**: 132 (Dec. 1931).
11. Physical Section, Training Division, Bureau of Navy Personnel, *Physical Fitness Manual for the U. S. Navy.* 1943, Ch. IV.
12. Rogers, Frederick Rand, *Physical Capacity Tests in the Administration of Physical Education.* New York: Bureau of Publications, Teachers College, Columbia University, 1926.
13. War Department, Field Manual, FM 21-20, *Physical Training.* Washington: Govt. Printing Office, 1951, Ch. 17.

(Submitted 12/21/54)

Comparative Analysis of Methods of Scoring Tests of Motor Learning

L. W. McCRAW
University of Texas
Austin, Texas

Abstract

Various methods have been used to measure the improvement made by individuals during repeated trials on a specific motor skill. This study compares eight such methods, some of which differ only in the selection of the final score. Considerable variations were found among the scores yielded by the methods, particularly when comparing improvements of individuals starting with different initial scores. Of the eight methods studied, the most acceptable appear to be those that relate gain to possible gain in the computation, while the least desirable are those that interpret gain in relation to the initial score.

A METHOD USED quite frequently to evaluate the ability of an individual to learn a specific motor skill is the so-called practice test. In using such tests, the objective is to determine the amount of improvement occurring during the performance of a definite number of trials on an activity involving the motor skill. The assumption is that if a fast learner and a slow learner having the same initial ability perform an act the same number of times, the fast learner will show more improvement during the performance than will the slow one.

It is generally agreed that an interpretation of the learning occurring under such circumstances is contingent on several factors. For example, Greene (11) stresses that, in order to compare the improvement of two individuals on the same learning test, every effort should be made to provide equal practice for all, equal motivation, constant behavior patterns, similar types of scores, equal units of measurements, and expression of improvement in terms of the difficulty of improvement. One of the hardest of these conditions to achieve is that of expressing the improvement in terms of difficulty.

Although many methods have been proposed to evaluate improvement, there is no indication that a satisfactory solution has been achieved. Rather it would appear that many of the methods yield entirely different results. It is the purpose of this study to examine certain of these methods, particularly with regard to the selection of scores to indicate initial and final status and to the interpretation of improvement made during repeated trials.

Discussion of the Problem

Various studies have indicated that there is little relationship or perhaps a low negative correlation between the initial status and improvement on repeated performances of a test (1, 8, 10). It is argued that it is harder for a person with the higher initial score to improve because he is much nearer the physiological limit of performance than is the person with the lower score.

The same difficulty arises even when two individuals start with identical scores, for as the individual improves over the initial score, it becomes increasingly more difficult for him to show additional gain. The measurement of improvement during a practice test then should be made in terms of where the individual is at the beginning of the practice period and of the relative difficulty of improvement.

The problem of interpreting the improvement made from an initial to a final status is not confined to learning tests but would be encountered in many experimental and practical situations in the use of pre-tests and post-tests to measure learning. Of those methods that have been used to express the improvement made under such circumstances, perhaps the simplest and most easily understood is that of computing the actual difference in raw scores by subtracting the initial score from the final score. Some advocate that raw scores should be converted to standard scores before obtaining such differences (6, 7). A most popular method (14) consists of dividing the difference in either raw scores or standard scores by the corresponding initial scores to obtain a per cent of improvement. It has been advocated by many investigators that this procedure penalizes unduly these individuals with high initial scores, and it has been suggested that the improvement should be divided by the possible gain rather than the initial score (3, 4, 14). The possible gain in this instance is computed by subtracting the initial score from the highest possible score that can be made on the test. Other methods of expressing improvement include an increased increment method (2, 5, 13) wherein more credit is given for performance as one progresses up the scale, and the absolute zero scale (11, 12), which is constructed by arbitrarily assigning zero to the highest negative standard representing the lowest raw score and then placing all other standard scores on the scale according to their distances from this lowest score. Although there are many similarities among these methods, there is evidence that several of them yield quite different results (14).

Another problem presented in the use of successive trials as a measure of learning is that of determining an initial score and a final score, particularly the latter. In most instances, the initial score consists of the first N number of trials at the beginning of the learning period. However, quite often, as a result of warming-up or familiarization, individuals exhibit a marked increase in scores immediately after the first two or three trials. It is usually recommended that these "warm-up" trials be discarded and not used as part of the initial score (9). Although there is no general agreement regarding the number of trials to be used, the initial score should include enough trials to insure its being a valid measure of the individual's initial status yet not so many as to involve improvement within itself.

It would seem logical that the final score should consist of a corresponding number of trials at the end of the practice period. Such procedure would be based on the assumption that there is improvement on each of the successive

trials in the test so that the score on the final trials will be greater than that on the initial trials. In actual practice, however, the scores are not always distributed in this manner. In many instances the learning curves reveal that many individuals achieve their maximum efficiency before the final trials, perhaps near the middle of the test, and then drop off because of fatigue, disinterest, or some unknown reason. It would appear therefore that in selecting the final score, consideration should be given to the maximum performance no matter where it occurs after the initial score.

Subjects, Tests, and Procedures

The subjects used in this study consisted of 134 seventh grade boys in the University High School, Austin, Texas. The ages of these boys ranged from 11 to 14, with 90 per cent of them between 12 and 13 years of age.

The tests used were a Rope Skip test and a Mirror Target Toss test. The Rope Skip test was taken by 128 boys and the Mirror Target Toss by 134. Both tests were administered in the regular physical education class period by the physical education instructor,[1] assisted by students from the Department of Physical and Health Education of the University of Texas.

The Rope Skip test consisted of having the subject skip a rope as rapidly as possible for 10 seconds while turning it himself. The subject stood with the rope behind his heels and started it on the command "Go." When a skip was missed, he stepped over the rope and continued until the 10-sec. trial was completed. Any type of skip was permitted as long as the subject jumped clearly over the rope as it passed beneath his feet. The score for a trial was the number of legal skips completed during the 10-sec. period. The highest score made on any one trial was 20. For the administration of the Rope Skip test, the class was divided into ten squads of four persons each. The 10-sec. trials on the test were given simultaneously to one person from each of the squads. Individuals within each squad alternated on the trials so that no one took two trials in succession. Each subject was given 21 trials with approximately 2 minutes' rest between trials. The entire test of 21 trials was administered during one class period.

The Mirror Target Toss test involved the subject's bouncing a volleyball off a smooth, hard-surfaced wall at a concentric target lying on the floor while he watched the image of the target in a mirror. The target consisted of three concentric circles, with diameters of 1, 3, and 5 feet respectively. The target was painted on a mat with its center 11 feet from the wall. The subject stood 5 feet from the wall and 11 feet to the left and slightly in front of the target. The mirror was placed to the right of the target and against the wall with its surface at a 45° angle from the wall. Such position caused the image in the mirror to be rotated 90° from the true position of the target. The subject was thus able to observe the target in the mirror as he faced the wall,

[1]Acknowledgment is made to James W. Tolbert of the University High School for his assistance in devising and administering tests.

but a screen placed between him and the target prevented him from looking directly at it at any time during the test. The test was scored by awarding three, two, and one points for hitting the inner, middle, and outer circles respectively. Each trial consisted of three throws, making a possible score of nine for one trial. Fifty such trials were taken by each of four subjects during one class period. Each subject was allowed five trials in succession, and the four subjects alternated so that there was approximately 4 minutes of rest for each person between performances.

Methods of Scoring Tests

The two learning tests were scored by eight methods as listed in Table 2. With the exception of the Total Learning Score Method, the computation of scores required the selection of initial and final scores. Some of the methods differ only in the procedure for computing the final score, whereas the initial score was the same for all methods.

The initial score for the Rope Skip test was the sum of the 4th and 5th trials after discarding the first three. In comparing the scores on the various trials, a significant difference was found between the means of the 3rd and 4th trials. Since, as previously discussed, this increase in performance was probably due to familiarization or warming-up, the first three trials were discarded. Inasmuch as performance thereafter was consistent, it was felt that two trials were adequate to indicate initial status. No significant increase was found among the initial trials for the Mirror Target Toss test; therefore, no trials were discarded. The first four trials were used as the initial score for this test since it was about twice as long as the Rope Skip test.

The final score was determined by three different procedures. The first, the one most commonly used, was the sum of the last two trials for the Rope Skip test and the sum of the last four trials for the Mirror Target Toss, the number of trials in each instance being comparable to that used for the initial score. The second procedure consisted of selecting the highest two successive trials for the Rope Skip and the highest four successive trials for the Mirror Target Toss occurring at any point after the initial score. A third procedure was simply using a total of all trials after the initial score. Each of the methods used in this study is described below:

I. Total Learning Score Method: This method consists of adding the scores on all trials in the learning test, including the initial and final scores.

II. Difference in Raw Score Method No. 1: The difference in raw score was obtained by subtracting the initial score from the final score, consisting of the sum of the last "N" trials as follows:

(Sum of last "N" trials) minus (Sum of first "N" trials)

III. Difference in Raw Score Method No. 2: This method differs from No. 1 above only in the computation of the final score.

(Sum of highest "N" successive trials) minus (Sum of first "N" trials)

IV. Per Cent Gain of Possible Gain Method No. 1: In the 3-per-cent gain of possible gain methods, the learning scores were computed by dividing the

actual gain from the initial score to the final score by the possible gain from the initial to the highest possible score. The three methods differed in the selection of the final scores as indicated in the formulas.

$$\frac{\text{(Sum of last "N" trials) minus (Sum of first "N" trials)}}{\text{(Highest possible score on "N" trials) minus (Sum of first "N" trials)}}$$

V. Per Cent Gain of Possible Gain Method No. 2:

$$\frac{\text{(Sum of highest "N" successive trials) minus (Sum of first "N" trials)}}{\text{(Highest possible score on "N" trials) minus (Sum of first "N" trials)}}$$

VI. Per Cent Gain of Possible Gain Method No. 3:

$$\frac{\text{(Sum of all trials) minus (Sum of first "N" trials)}}{\text{(Highest possible score on all trials) minus (Sum of first "N" trials)}}$$

VII. Per Cent Gain of Initial Score Method No. 1: In the 2-per-cent gain of initial score methods, the scores were computed by dividing the gain from the initial to the final score by the initial score. Two different procedures of selecting the final score were used as indicated in the formulas.

$$\frac{\text{(Sum of last "N" trials) minus (Sum of first "N" trials)}}{\text{(Sum of first "N" trials)}}$$

VIII. Per Cent Gain of Initial Score Method No. 2:

$$\frac{\text{(Sum of highest "N" successive trials) minus (Sum of first "N" trials)}}{\text{(Sum of first "N" trials)}}$$

Findings

The analysis of the eight methods of scoring the two tests of motor learning used in this study is made by examining (1) the variability of scores yielded by the various methods, (2) the relationship among the methods, and (3) the validity of the different methods. For this analysis, the subjects were divided into three approximately equal groups on the basis of low, moderate, or high initial scores and into three groups on the basis of small, moderate, or large gain. The gain used for this subdivision is the gain in raw scores from the initial to the final score (II. Difference in Raw Score Method No. 1).

COMPARISON OF VARIABILITY

A comparison of the variability of scores yielded by the eight methods may be made by examining the coefficients of variation in Table 1. A perusal of this table reveals that there are marked differences in variability among most of the methods for both the total group and each of the subgroups. Likewise, substantial differences are noted within each method for the various subgroups.

With few exceptions the lowest coefficients are found for the Total Learning Score Method and the Per Cent Gain of Possible Gain Method No. 3, which involves the total score in the computation (Methods I and VI). The highest coefficients are found for the Per Cent Gain of Initial Score Methods (Methods VII and VIII). Very little difference is noted between the coefficients for the Difference in Raw Score Methods and the Per Cent Gain of Possible Gain Methods (Methods II, III, IV, and V). These conclusions are based on a comparison of methods with similar procedure of selecting the final score. When the different methods of selecting the final score are com-

TABLE 1

Coefficients of Variation on the Different Methods of Scoring Learning for the Various Subgroups Selected on the Basis of Initial Score and Amount of Improvement

Scoring Method	I TLS		II DRS No. 1		III DRS No. 2		IV PPG No. 1		V PPG No. 2		VI PPG No. 3		VII PIS No. 1		VIII PIS No. 2	
	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT
Low Initial Score	30	57	58	64	41	50	61	63	43	49	25	58	64	81	45	70
Moderate Initial Score	16	33	69	62	41	36	70	61	43	34	17	33	73	78	42	56
High Initial Score	15	26	98	100	47	51	102	102	54	46	16	28	98	104	46	65
Small Gain	31	61	107	129	41	52	106	128	50	61	32	62	118	282	58	167
Moderate Gain	31	33	56	68	11	15	54	68	27	17	34	33	86	185	48	123
Large Gain	22	21	40	33	16	13	45	32	30	12	23	19	50	88	30	85
Total Group	30	41	75	84	45	54	81	77	71	45	31	41	91	155	57	139

TLS—Total Learning Score; DRS—Difference in Raw Score; PPG—Per Cent Gain of Possible Gain; PIS—Per Cent Gain of Initial Score; RS—Rope Skip Test; MTT—Mirror Target Toss Test.

pared, it is found that in general the larger coefficients were obtained in those methods using the sum of the last N trials as the final score (Methods II, IV, and VII) rather than the sum of the highest N successive trials (Methods III, V, and VIII). No consistent pattern is noted among the various subgroups; however, by far the largest coefficients are found for the groups with small gain and the lowest for the ones with large gain.

RELATIONSHIP AMONG METHODS

In order to determine the relationship among methods, intercorrelations were computed for the total group and the subgroups. Correlations for the subgroups with low, moderate, and high initial scores are shown in Table 2, and those for the subgroups with small, moderate, and large gains in Table 3. Correlations for the total group are included in both tables to facilitate comparisons.

An examination of the correlations in these two tables reveals that there is practically no difference between the Total Learning Score Method and the Per Cent Gain of Possible Gain Method No. 3 (Methods I vs. VI). All correlations are .93 or above with most being .99. Such relationship is not surprising since the total score is used as the final score in the computation of this Per Cent Gain of Possible Gain Method. In view of this close relationship, further comparison will be made only on the basis of the Total Learning Score Method. Very high correlations are also found between the Difference in Raw Score Methods and the Per Cent Gain of Possible Gain Methods (Methods II vs. IV and III vs. V). With the exception of those for the subgroups with moderate and large gains, the correlations are all above .80 with most being above .90. The most pronounced disagreement among the methods of scoring exists between the Total Learning Score Method and each of the two Per Cent Gain of Initial Score Methods (Methods I vs. VII and I vs. VIII). Very low positive or moderately high negative correlations are found for these combinations.

The relationship among the different procedures of selecting the final score may be ascertained by examining the correlations between the two methods for the respective systems of computing gain (Methods II vs. III, IV vs. V, and VII vs. VIII). Although there is no consistent pattern with the correlations ranging from .90 to —.18, it is apparent that there is only moderate or slight relationship among the scores yielded by these two procedures.

There seems to be no indication that higher or lower correlations exist for any of the three subgroups in Table 2 or for any of the three in Table 3. However, when the correlations in Table 2 are compared with the corresponding correlations in Table 3, it is readily seen that in most instances much lower positive or higher negative correlations are found in the latter table. This would suggest that where differences among methods do exist, they are more pronounced for the subgroups with small, moderate, or large gains. Such marked deviations are probably due to the fact that individuals within each group start from varying initial scores rather than that they have different gains.

TABLE 2

*Coefficients of Correlation for Subgroups with Low, Moderate, and High Initial Scores**

Scoring Method	I	II	III	IV	V	VI	VII	VIII
I. Total Learning Score Method		.43	.51	.53	.63	.95	.05	.01
		.71	.70	.72	.76	.99	.63	.56
		.40	.62	.52	.82	.99	.31	.42
		.07	.30	.31	.68	.99	−.32	−.36
II. Difference in Raw Score Method No. 1	.79		.69	.99	.65	.39	.89	.56
	.63		.75	.99	.74	.72	.99	.72
	.31		.74	.96	.64	.41	.99	.74
	.27		.71	.89	.50	.08	.83	.64
III. Difference in Raw Score Method No. 2	.91	.70		.67	.98	.55	.53	.82
	.88	.53		.75	.99	.71	.74	.97
	.27	.40		.73	.92	.63	.72	.96
	.38	.73		.68	.83	.30	.51	.41
IV. Per Cent Gain of Possible Gain Method No. 1	.80	.99	.71		.68	.49	.82	.50
	.63	.99	.63		.75	.76	.97	.70
	.46	.95	.29		.71	.53	.91	.66
	.40	.96	.65		.65	.32	.62	.43
V. Per Cent Gain of Possible Gain Method No. 2	.92	.70	.99	.71		.57	.45	.70
	.88	.60	.98	.61		.77	.71	.63
	.50	.37	.89	.37		.83	.56	.78
	.62	.60	.91	.61		.68	.18	.31
VI. Per Cent Gain of Possible Gain Method No. 3	.99	.78	.92	.79	.92		.01	−.05
	.96	.65	.85	.65	.87		.64	.57
	.98	.32	.30	.46	.52		.32	.43
	.99	.30	.40	.42	.63		−.32	−.36
VII. Per Cent Gain of Initial Score Method No. 1	.57	.83	.59	.79	.55	.56		.70
	.54	.89	.63	.83	.51	.53		.75
	.15	.96	.48	.85	.37	.18		.76
	−.06	.69	.55	.55	.34	−.03		.80
VIII. Per Cent Gain of Initial Score Method No. 2	.62	.55	.80	.54	.75	.63	.68	
	.65	.62	.83	.54	.70	.60	.81	
	.15	.45	.92	.29	.74	.15	.59	
	−.08	.54	.65	.40	.42	−.04	.83	

*Entries for each combination are for subgroups with low, moderate, and high initial scores and for the total group respectively.

Correlations in the upper portion of the table are for the Rope Skip Test while those in the lower portion are for the Mirror Target Toss Test.

TABLE 3

*Coefficients of Correlation for Subgroups with Small, Moderate, and Large Gains**

Scoring Method	I	II	III	IV	V	VI	VII	VIII
I. Total Learning Score		−.14	.21	.04	.63	.99	−.42	−.49
		−.34	−.13	.09	.89	.99	−.65	−.84
Method		.02	.19	.39	.83	.99	−.41	−.70
		.07	.30	.31	.68	.99	−.32	−.36
II. Difference	.04		.33	.97	.11	−.13	.89	.44
in Raw Score	−.21		.33	.88	−.21	−.33	.84	.42
Method	.32		.53	.85	.11	.03	.84	.39
No. 1	.27		.71	.89	.50	.08	.83	.64
III. Difference in	.34	.16		.36	.84	.22	.28	.65
Raw Score	−.13	.09		.32	.25	−.14	.23	.42
Method	.16	.27		.49	.55	.19	.39	.46
No. 2	.38	.73		.68	.83	.30	.51	.41
IV. Per Cent	.32	.90	.18		.24	.05	.78	.32
Gain of Pos-	−.01	.96	.04		.20	.09	.71	.09
sible Gain	.45	.96	.17		.48	.42	.49	.01
Method	.40	.96	.65		.65	.32	.62	.43
No. 1								
V. Per Cent	.68	.13	.81	.33		.64	.38	.24
Gain of Pos-	.70	−.32	.44	−.18		.86	−.49	−.59
sible Gain	.54	.12	.71	.20		.83	.12	−.35
Method	.62	.60	.91	.61		.68	.18	.31
No. 2								
VI. Per Cent	.99	.04	.35	.32	.69		−.41	−.48
Gain of Pos-	.99	−.20	−.13	.00	.68		−.65	−.85
sible Gain	.99	.38	.08	.52	.45		−.41	−.70
Method	.99	.30	.40	.42	.63		−.32	−.36
No. 3								
VII. Per Cent	−.28	.70	.11	.41	−.08	−.29		.67
Gain of Ini-	−.54	.59	.03	.41	−.50	−.54		.75
tial Score	−.23	.52	.60	.32	.10	−.19		.75
Method	−.06	.69	.55	.55	.34	−.03		.80
No. 1								
VIII. Per Cent	−.45	.33	.27	.11	−.04	−.44	.64	
Gain of Ini-	−.67	.29	.26	.11	−.48	−.67	.44	
tial Score	−.39	.20	.67	.01	.09	−.37	.90	
Method	−.08	.54	.65	.40	.42	−.04	.83	
No. 2								

*Entries for each combination are for subgroups with small, moderate, and large gains and for the total group respectively.

Correlations in the upper portion of the table are for the Rope Skip Test while those in the lower portion are for the Mirror Target Toss Test.

VALIDITY OF METHODS OF SCORING

An indication of the validity of the scoring methods may be obtained by comparing the scores yielded by the different methods for the various subgroups. This comparison is made on the basis of mean T-scores computed for each subgroup for each method. The T-score scales used for these computations were based on the total distribution for each method of scoring.

In Table 4 are presented the mean T-scores for the subgroups arranged on the basis of initial score. For example, the first three subgroups all start with a low initial score but the three have small, moderate, and large gains respectively; the second three subgroups all start with a moderate initial score and have small, moderate, and large gains; while the last three start with a high initial score. To be valid a scoring method should yield increasingly larger scores within each of the three subgroups starting from the same initial score, i.e., the score for the subgroup with Low Initial Score and Moderate Gain should be larger than that for the one with Low Initial Score and Small Gain and so on. While there are differences among the methods with regard to the amount of increase among subgroups, all methods meet this criterion.

In Table 5 the same mean T-scores have been arranged on the basis of improvement rather than initial score. The first three, for example, all have small gain but start with low, moderate, and high initial scores; the second three have moderate gain from varying initial scores; and the third three have large gain. Again, to be valid a scoring method should yield increasingly larger scores within each set of subgroups. This is based on the contention that it becomes more difficult to exhibit the same improvement as one approaches his physiological limit. Thus the subgroup with a Small Gain over a High Initial Score should have a larger learning score than those with Small Gain over a Low or Moderate Initial Score. When evaluated on the basis of this criterion, only three methods are valid; namely, I. Total Learning Score Method; V. Per Cent Gain of Possible Gain Method No. 2; and VI. Per Cent Gain of Possible Gain Method No. 3. It will be recalled that Method V uses for the final score the highest N successive trials no matter where they occur after the initial score and Method VI uses the total score as the final score. The two Per Cent Gain of Initial Score Methods (Methods VII and VIII) are the least desirable of all methods; in fact, they appear to be wholly invalid when comparing individuals starting from different initial scores.

Summary and Conclusions

The purpose of this study has been to compare eight methods of scoring improvement on learning tests in which individuals were given a definite number of repeated trials on a specific motor skill. Some of the methods differ only in the selection of the final score—the sum of the *last* two to four trials in one instance and in the other the sum of the *highest* two to four successive trials occurring after the initial score. The initial score was the sum

TABLE 4

Mean T-Scores on the Different Methods of Scoring Learning for the Various Subgroups Arranged on the Basis of Initial Score

Subgroup	I TLS		II DRS 1		III DRS 2		IV PPG 1		V PPG 2		VI PPG 3		VII PIS 1		VIII PIS 2	
	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT
Low Initial Score Small Gain	37	30	45	44	40	38	44	43	44	35	38	30	48	51	46	49
Low Initial Score Moderate Gain	39	39	54	50	51	51	50	48	51	47	39	39	61	57	62	57
Low Initial Score Large Gain	45	51	59	61	60	63	55	59	58	59	45	52	61	66	64	68
Moderate Initial Score Small Gain	44	38	43	40	39	39	44	39	45	37	44	37	43	44	40	44
Moderate Initial Score Moderate Gain	48	49	48	51	50	52	47	52	54	52	48	49	47	46	49	46
Moderate Initial Score Large Gain	54	58	60	59	63	60	59	59	65	60	55	57	55	49	57	49
High Initial Score Small Gain	56	54	39	42	40	39	41	44	49	42	56	54	40	44	39	43
High Initial Score Moderate Gain	62	56	48	45	50	49	51	47	61	54	62	56	44	45	43	44
High Initial Score Large Gain	66	58	55	53	60	47	61	56	74	61	66	59	48	46	48	46

TLS—Total Learning Score; DRS—Difference in Raw Score; PPG—Per Cent of Possible Gain; PIS—Per Cent of Initial Score; RS—Rope Skip Test; MTT—Mirror Target Toss Test.

TABLE 5

Mean T-Scores on the Different Methods of Scoring Learning for the Various Subgroups Arranged on the Basis of Improvement

Subgroup	I TLS		II DRS 1		III DRS 2		IV PPG 1		V PPG 2		VI PPG 3		VII PIS 1		VIII PIS 2	
	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT	RS	MTT
Small Gain Low Initial Score	37	30	45	44	40	38	44	43	44	35	38	30	48	51	46	49
Small Gain Moderate Initial Score	44	38	43	40	39	39	44	39	45	37	44	37	43	44	40	44
Small Gain High Initial Score	56	54	39	42	40	39	41	44	42	42	56	54	40	44	39	43
Moderate Gain Low Initial Score	39	39	54	50	51	51	50	48	51	47	39	39	61	57	62	57
Moderate Gain Moderate Initial Score	48	49	48	51	50	52	47	52	54	52	48	49	47	46	49	46
Moderate Gain High Initial Score	62	56	48	45	50	49	51	47	61	54	62	56	44	45	43	44
Large Gain Low Initial Score	45	51	59	61	60	63	55	59	58	59	45	52	61	66	64	68
Large Gain Moderate Initial Score	54	58	60	59	63	60	59	59	65	60	55	57	55	49	57	49
Large Gain High Initial Score	66	58	55	53	60	47	61	56	74	61	66	59	48	46	48	46

TLS—Total Learning Score; DRS—Difference in Raw Score; PPG—Per Cent of Possible Gain; PIS—Per Cent of Initial Score; RS—Rope Skip Test; MTT—Mirror Target Toss Test.

of a comparable number of trials at the beginning of the practice period. Four different procedures were used in the computation of improvement as follows: the summation of scores of all trials on the test, subtracting the initial score from the final score, dividing the gain from the initial to the final score by the initial score, and dividing this gain by the possible gain after the initial score.

The learning tests were two activities in which the subjects had little or no experience. A Rope Skip test consisted of determining the number of times an individual could skip a rope while turning it himself during each of 21 10-sec. trials. A Mirror Target Toss test involved bouncing a volleyball off a wall at a concentric target lying on the floor while watching the image of the target in a mirror. Fifty trials of three throws each were allowed.

Comparison of the eight methods was made on the basis of coefficients of variation, Product-moment correlations, and mean T-scores computed for the total group and for subgroups with varying initial scores and gain. The following conclusions seem warranted:

1. There are marked differences in variability of scores yielded by the different methods suggesting that some of the methods allow unusually high or low scores for certain individuals.
2. In general, there is rather low or negative relationship among most of the methods of scoring, particularly when comparing improvements of individuals starting with widely different initial scores.
3. Of those methods used in this study, the Total Learning and the three Per Cent Gain of Possible Gain Methods seem to be the most valid methods when used to compare individuals with different initial scores.
4. The two Per Cent Gain of Initial Score Methods seem to be the least desirable and should not be used to measure improvement on learning tests.

REFERENCES

1. Anastasi, Anne. Practice and Variability. *Psychological Monographs*, **45**: 1-55, 1934.
2. Bovard, John F., Frederick W. Cozens, and E. Patricia Hagman. *Tests and Measurements in Physical Education*. Philadelphia: W. B. Saunders Co., 1949.
3. Brace, David K. Studies in motor learning of gross bodily motor skills. *Research Quarterly*, **17**: 242-253, 1946.
4. ———. Studies in the rate of learning gross bodily motor skills. *Research Quarterly*, **12**: 181-185, 1941.
5. Cozens, Frederick W. A curve for devising scoring tables in physical education. *Research Quarterly*, **2**: 67-75, 1931.
6. Cureton, Thomas K. The physiology of fitness. *Scholastic Coach*, **13**: 14-20, 1944.
7. ———, Warren J. Huffman, Lyle Welser, Ramon W. Kireilis, and Darrell E. Latham. *Endurance of Young Men*. Monograph of the Society for Research in Child Development, 1945, 1.
8. Gates, Georgina Strickland. Individual differences as affected by practice. *Archives of Psychology*, **8**: 1-74, 1922.
9. Gire, Eugenia, and Anna Espenschade. The relationship between measures of motor educability and the learning of specific motor skills. *Research Quarterly*, **13**: 43-56, 1942.

10. GOODENOUGH, FLORENCE L., and CLARA R. BRIAN. Certain factors underlying the acquisition of motor skill by pre-school children. *Experimental Psychology*, **12**: 127-155, Apr., 1929.
11. GREENE, EDWARD B. *Measurement of Human Behavior*. New York: The Odyssey Press, 1941.
12. ———. Practice effects on various types of standard tests. *American J. of Psychology*, **49**: 67-75, 1937.
13. MCCLOY, CHARLES H. *The Measurement of Athletic Power*. New York: A. S. Barnes and Co., 1932.
14. MCCRAW, L. W. A comparison of methods of measuring improvement. *Research Quarterly*, **22**: 191-200, 1951.

(Submitted 11/1/54)

Cardiovascular Measurements on College-Age Smokers and Non-Smokers

NICHOLAS PETER MOUTIS
The Phillips Exeter Academy
Exeter, New Hampshire

Abstract

Thirty-six male college students were measured in six cardiovascular factors as recorded by the Cameron Heartometer. Seventeen of these subjects were smokers and 19 were non-smokers. The purpose of this investigation was to find what residual, as opposed to acute, effects, smoking has on young men. Each subject in both groups was measured in six cardiovascular factors, nine times successively, during his testing period. The smokers abstained from smoking on the day of testing.

The residual effects of smoking in young men, if there be any, cannot be demonstrated in a statistically significant manner in certain cardiovascular factor measurements in the quiet state.

RECENTLY, THE controversial subject of the smoking of tobacco, and of its harmful effects upon the health of those who smoke, has been pursued with exceptional vigor. The bulk of this research and experimentation has concerned itself with the acute effects of smoking insofar as physical performance and physiologic changes are concerned. Little, if any, attention has been focused upon the possible residual effects smoking may have on the physiologic constitution, and yet we hear almost daily, especially in athletic circles, of the deleterious results of tobacco-smoking on the heart and circulatory system.

Review of the Literature

Recent studies have emphasized, as mentioned, the acute effects of cigarette smoking. Wilgoose (11) and Kay and Karpovich (5) illustrate the difficulty of standardizing test procedures. The former tested and found recovery rate was greater after a non-smoking bout as opposed to results after a period of smoking. The latter states that one cigarette by an habitual smoker had no effect on recovery. It seems that habituation is a moot point especially where the so-called tobacco-sensitive individual is concerned. In a study that provoked some controversy, Levy et al (6) found no significant changes after smoking in normal persons or persons with cardiac disease. In this respect, Stewart (10) found that physiologic changes induced by tobacco smoking were less marked in elderly patients than they were in young, normal subjects.

Investigations of both the acute and residual effects of smoking have shown that the most consistent index of the physiological effects of smoking was the pulse rate increase. Parker (8) found this so in 19 subjects between the ages of 19 and 25 years. Autio et al (1) and Schilpp (9) also found this to be

true and the latter includes mean heart rates for smokers at 72.7 beats per minute and for non-smokers at 66.2 beats per minute.

That this rate increase persists in smokers even in violent exercise is reiterated by Haggard (4), who states that rates as high as 200 beats per minute or even higher occur in untrained men (especially those who smoke habitually) during vigorous exertion.

It is interesting to note that in a study of 2,577 naval enlistees reported by Ludwick and Massler (7), 75 per cent smoked and 90 per cent of these smokers used more than half a pack daily. This may answer in part how much young smokers, in this case between 17 and 21 years of age, smoke. In reviewing the literature, the writer could not find in any one study, an investigation of the residual effects of smoking on more than one or two physiologic functions.

Thus, this investigation attempted to determine what residual differences in a number of certain cardiovascular measures exist between young men who smoke and others who do not smoke.

Methodology

Thirty-six male college students were used in the investigation. The subjects were selected at random from the University of Illinois Reserve Officers Training Corps in the spring of 1953. Their co-operation was enchanced by the fact that they were paid at a standard rate for the use of their time. Seventeen were smokers and 19 were non-smokers. Their ages ranged from 17 to 20 years.

An orientation period for all the subjects was conducted individually. During this period the subject familiarized himself with the various pieces of apparatus to be used, and a brief explanation was given regarding the purpose of the study. A short practice run followed on all the equipment in order to eliminate any apprehension and to enable each man to become accustomed to the procedure that was to follow. An air-conditioned room in the Climatic Research Laboratory was used and kept comfortable at 72° F. and 50 per cent humidity during the testing.

Each subject was tested twice, at approximately the same time of day on both occasions, usually one week apart. The time of the testing was at least two hours after the ingestion of meals, and all the men in the smoking group were instructed to abstain from smoking on the day they presented themselves for testing. The reliability re-tests attempted, as far as controllably possible, to adhere to the same pattern as the first test conditions.

On the testing day, the subject arrived and was seated in an ordinary wooden chair and permitted to relax for ten minutes prior to the beginning of the experiment. During this period he was being prepared for the testing, and all the apparatus was placed in position. At the close of this ten-minute interval a basal measurement was taken with the Cameron Heartometer.

Systolic blood pressure, diastolic blood pressure, pulse pressure, pulse rate, systolic amplitude, and diastolic amplitude measurements were recorded.

Immediately following this initial measurement, the subject read a passage which took approximately two minutes. As soon as he had finished the reading, another Heartometer record was taken of all the measurements stated previously. The completion of this procedure varied between 30 seconds and one minute. At the end of this second recording, the subject read another passage. The same procedure followed after this reading. In all, eight passages for reading were administered, giving a total of nine records of all the physiological measures, one basal and eight sets of measurements after eight reading passages. The administration of the complete test, from basal through the eight passages, took approximately 30 minutes. The same procedure was used in the re-test, which was given one week later.

APPARATUS

The Cameron Heartometer was the instrument used for recording the cardiovascular measurements selected for the investigation. It is a direct-writing instrument powered electrically, and details of the mechanism may be found in a publication by Cureton (2) from the Laboratory of Physical Fitness at the University of Illinois.

EXPERIMENTAL VARIABLES

Measurements were taken and recorded on:

1. Pulse Rate
2. Systolic Blood Pressure
3. Diastolic Blood Pressure
4. Pulse Pressure
5. Systolic Amplitude
6. Diastolic Amplitude

STATISTICAL TREATMENT

Averages of basal measurements were computed for the six factors of the smoking group versus the non-smoking group with corresponding t-values. These scores are shown in Table 1. Table 2 shows the averages of the six factors after each reading passage with corresponding t-values.

The resulting t-values are compared both at the .01 and .05 per cent level of confidence for 34 cases. In some instances, there are only 33 or 32 cases available. This makes the value of t somewhat higher at both the .01- and .05-per-cent-level of confidence.

Reference to Garret (3) was made to determine the significance of the t-value obtained at both levels of confidence.

Discussion of Results

Table 1 shows the result of the basal measures, taken on each of the six factors, with corresponding t-values.

In the smoking group, mean pulse rates, systolic blood pressures, pulse pressures and systolic amplitudes are higher than in the non-smoking group.

TABLE 1

Mean-Basal Measures of Smokers and Non-Smokers on Six Cardiovascular Factors

Factors	Smokers (17)	Non-Smokers (19)	M.-diff.	t
Pulse Rate (beats/min.)	76.28	73.78	3.04	.85
Systolic B. P. (mm/Hg)	123.00	121.63	1.37	.55
Diastolic B. P. (mm/Hg)	73.82	76.05	2.23	1.03
Pulse Pressure (mm/Hg)	48.94	44.78	4.16	.86
Systolic Amplitude (mm)	10.1	9.7	.4	.47
Diastolic Amplitude (mm)	4.3	4.49	.19	.43

However, these mean differences were not significant at the .05-per-cent-level of confidence.

In Table 2, each figure represents 17 cases in the smoking group and 19 cases in the non-smoking group. Each t-ratio in Table 2 represents 34 cases, except where noted by asterisk and in that case represents either 33 or 32 cases. It is evident again that few of the scores are statistically significant. Only five t-ratios are significant at the .05-per-cent-level of confidence.

The scores from Table 3 are averages computed from the results of the eight reading passages. The largest mean difference between the two groups after the reading passages is in the pulse rate. Also, it is obvious that the pulse rate t-value is highest in Table 3. This is in accord with Parker (8), as previously cited. It may be assumed that these reading passages produced slight effect on the six factor measurements. They seem even to exhibit a quieting influence on these dynamic factors. This was evidenced by the fact that most scores were lowered as the reading progressed from passage one to passage eight. Of some note is the observation that systolic blood pressure tended to decrease and increase systematically in both smokers and non-smokers during the reading of the eight passages. A drop in systolic blood pressure of eight mm/Hg and seven mm/Hg respectively in smokers and non-smokers took place during these readings. At no time, however, do the systolic blood pressure t-values approach significance at the .05-per-cent-level of confidence.

Generally, larger changes occurred in the smoking group than in the non-smoking group when a comparison is made of the basal scores in Table 1 and the mean scores of the eight readings found in Table 3. This may lend some credence to the inference often made, not based on fact, that smokers tend to be a nervous, jittery group. On the other hand, smokers as a group may be better able to relax in a quiet situation than can non-smokers, as evidenced by the size of the decrease in their scores when comparing Table 1 with Table 2. A specific citation regarding decreases in certain factors may be helpful. The pulse pressure diminution is used as an example. It is important to note that at the end of the eighth reading passage the decrease in pulse pressure was almost ten times as large in the smokers as it was in the

TABLE 2

Means and t-Values of Six Factors on Smokers vs Non-Smokers Taken After Eight Reading Passages

Reading Passages	1	2	3	4	5	6	7	8
	Smokers (17)							
Pulse Rate (beats/min)	79.76	76.23	78.11	77.70	76.70	78.47	76.66	75.46
Systolic B. P. (mm/Hg)	119.88	117.58	117.64	117.35	111.41	113.23	111.60	111.40
Diastolic B. P. (mm/Hg)	72.17	70.82	62.29	69.82	68.82	70.73	73.30	73.40
Pulse Pressure (mm/Hg)	47.70	47.05	48.35	47.52	42.70	42.35	38.20	37.86
Systolic Amplitude (mm)	9.6	9.7	9.1	9.6	9.5	9.9	9.3	9.8
Diastolic Amplitude (mm)	4.01	3.92	3.61	3.85	4.00	3.65	3.69	3.74
	Non-Smokers (19)							
Pulse Rate (beats/min)	72.22	73.68	73.89	74.73	74.10	73.68	74.73	73.05
Systolic B. P. (mm/Hg)	120.84	121.89	119.89	117.89	97.89	116.68	115.21	113.52
Diastolic B. P. (mm/Hg)	78.26	73.21	72.26	71.57	70.63	70.15	71.42	71.26
Pulse Pressure (mm/Hg)	46.36	48.63	47.47	46.31	44.36	46.89	43.78	43.42
Systolic Amplitude (mm)	9.6	9.0	9.8	9.5	9.5	9.1	9.9	9.5
Diastolic Amplitude (mm)	4.25	3.82	4.39	4.10	4.33	4.36	4.41	4.03
	t-Values							
Pulse Rate (beats/min)	2.32*	.91	1.66	.88	.63	1.46	.65**	.72**
Systolic B. P. (mm/Hg)	.26	.99	.52	0	2.39	.90	1.36**	.57**
Diastolic B. P. (mm/Hg)	2.31	1.11	1.26	.70	.54	0	.52**	.57**
Pulse Pressure (mm/Hg)	.22	.20	.21	.20	.35	.87	1.02**	1.18**
Systolic Amplitude (mm)	0	.86	.85	.11	0	.90**	.61**	.33**
Diastolic Amplitude (mm)	.75*	.24	1.91	.43	.70	1.47	1.36**	.77**

* = 33 cases.
** = 32 cases.

TABLE 3

Means, Mean Differences and t-Values of Six Cardiovascular Factors of Smokers vs. Non-Smokers Following Eight Reading Passages

Reading Passages	Smokers (17)	Non-Smokers (19)	M.-diff.	t
Pulse Rate (beats/min)	78.38	73.76	4.62	1.15
Systolic B. P. (mm/Hg)	115.00	115.60	.60	.87
Diastolic B. P. (mm/Hg)	71.05	72.34	1.29	.87
Pulse Pressure (mm/Hg)	43.96	45.90	1.94	.53
Systolic Amplitude (mm)	9.56	9.48	.08	.45
Diastolic Amplitude (mm)	3.80	4.21	.41	.95

non-smokers. This would indicate, for pulse pressure at least, that the smokers were more affected by the eight reading passages than were the non-smokers. As to this point, Cureton et al. (2) state that pulse pressure will diminish with fatigue but that this item is quite unreliable in predicting endurance.

The systolic and diastolic amplitudes were not consistent measures in this investigation. The smokers displayed higher systolic amplitudes and the non-smokers showed higher diastolic amplitudes. The mean differences were slight and insignificant.

Summary and Conclusions

Using the Cameron Heartometer, 36 male college students were tested in six cardiovascular factors. Heartograms were taken on each individual in a sitting-basal position and immediately after each of eight reading passages of approximately two minutes' duration.

Seventeen smokers and 19 non-smokers made up the group. The investigation was concerned with determining what residual differences, if any, exist between smokers and non-smokers in six measures of cardiovascular functioning on a day on which no smoking is permitted.

From the results obtained, it may be concluded:

1. The smoking group showed higher, though not significant, mean pulse rates, systolic blood pressures, pulse pressures, and systolic amplitudes in the sitting-basal position.

2. The pulse rate was the most consistently higher item in the smokers throughout the test. This result is in full accord with Parker (8).

3. Generally, the largest mean difference among the cardiovascular factors was found in the pulse-rate. This difference was 3.04 beats/minute in the sitting-basal position and 4.62 beats/minute after the eight reading passages, with the smoking group having the higher rate.

4. Pulse pressure decreased 5 mm/Hg for the smokers throughout the test, whereas the non-smokers experienced a rise. This may indicate a fatigue effect, Cureton et al. (2), in the smokers.

5. The smokers elicited larger changes in most of the six cardiovascular factors than did the non-smokers during the testing.

6. The results tend to show that the reading passages had a more quieting effect on the non-smokers than on the smokers.

7. The residual effects of smoking in young men, if there are any, could not be demonstrated in a statistically significant manner in the selected cardiovascular factors.

REFERENCES

1. Autio, L., O. Eranko, and E. Jalavisto. Vasomotor Reactions in Valsalva's Experiment. *Acta Physiologica Scandinavia*, **17**: 130-149, 1949.
2. Cureton, T. K., et al. *Physical Fitness Appraisal and Guidance*. St. Louis: C. V. Mosby Co., 1947.
3. Garret, H. K. *Statistics in Education and Psychology*. New York: Longmans, Green and Co., 1930.
4. Haggard, H. W. *The Science of Health and Disease*. New York: Harper and Bros., 1938.
5. Kay, H. W., and P. V. Karpovich. Effect of Smoking Upon Recuperation from Local Muscular Fatigue. *Research Quarterly*, **20**: 250-256, 1949.
6. Levy, R. L., et al. Effects of Smoking Cigarettes on the Heart. *American Medical Association*, **135** 417-422, 1947.
7. Ludwick, W., and M. Massler. Relation of Dental Caries Experience and Gingivitis to Cigarette Smoking in Males Seventeen to Twenty-One Years Old. *Journal of Dental Research*, **31**: 319-322, 1952.
8. Parker, P. A. Acute Effects of Smoking on Physical Endurance and Resting Circulation. *Research Quarterly*, **25**: 210-217, 1954.
9. Schilpp, R. W. A Mathematical Description of the Heart Rate Curve of Response to Exercise with Some Observations on the Effects of Smoking. *Research Quarterly*, **22**: 439-445, 1951.
10. Stewart, H. J., et al. The Effect of Smoking Cigarettes on the Peripheral Blood Flow in Subjects in the Older Age Group with Coronary Arteriosclerosis and Hypertension. *American Heart Journal*, **30**: 541-550, 1945.
11. Wilgoose, C. E. Tobacco Smoking, Strength, and Muscular Endurance. *Research Quarterly*, **18**: 219-225, 1947.

(Submitted 2/9/55)

Health and Physical Education Program for Boys in Ohio High Schools[1]

MAJOR LAURENCE E. OWENS, USAF
Indiana University
Bloomington, Indiana

Abstract

This study was undertaken to determine the status of the health and physical education program for boys in 101 Ohio secondary schools and to determine the relationship of school size, community size, accreditation, geographic area, type of school district, and consolidation to this status. According to the results of this study, accreditation and school size have considerably more influence upon the quality of the program than any other factor studied. Schools enrolling 500-999 students have a significantly better health and physical education program than schools enrolling less than 500 students. Enrollment above 1,000 tends to have a slight negative influence upon the quality of the program.

THE PURPOSE OF this study is to determine the status of the boys' health and physical education program in selected Ohio public secondary schools and to analyze the relationship of selected factors to this status. The study is limited to the public first grade high schools as defined and approved by Ohio State Department of Education and listed in the Educational Directory (5) for the school year 1951-52.

Independent variables considered are size of school enrollment, size of community, accreditation by the North Central Association of Colleges and Secondary Schools, geographic area,[2] type of school district, and consolidation.

Personal visitation and interview with application of the *Health and Physical Education Score Card No. II* (3) during the 1952-53 school year is the method used in gathering data from the schools.

A pilot study of nine Indiana high schools was made to acquire skill in the application of the score card prior to the evaluation of the Ohio sample.

Procedure

There were 1,040 schools retained for consideration in the selection of schools for this study. It did not appear feasible to evaluate all schools, so it was decided to select a controlled sample (4, *p. 5*) of 100 schools. The assumption was made that 100 schools were an ample number for the pur-

[1]This study was made in partial fulfillment of the requirements for the degree of Doctor of Physical Education in the School of Health, Physical Education, and Recreation, Indiana University, 1955, under the direction of Dr. Karl W. Bookwalter.

[2]The 1950 Census divides the State of Ohio into four geographic areas: northwestern, northeastern, southeastern, and southwestern.

pose of this study. A controlled sample was used because it was desired to make the sample representative with respect to size of school and size of community.

Basic categorizations suggested in the *Research Bulletin* of the National Education Association (8) for the stratification of each school into school-size—town-size cells were gleaned. Valid enrollment (5) and population (7, 10) data were secured for each of the 1,040 Ohio schools. Suggested enrollment and population categorizations were adjusted to fit the distribution found in Ohio.

Each school was assigned a small card. An identifying number was written on the card and the card was placed in the appropriate cell. The cards were then turned face down, shuffled, and renumbered consecutively from one to N in each cell.

A sample of 100 schools is 9.6 per cent of the total schools retained for selection. The number of schools in each cell was multiplied by .096 and the product rounded to the nearest whole number. This number represented the predetermined number of schools to be taken from each cell.

Table 1 shows the basic categories of enrollment and population used and the proportionate number of schools to be selected.

An exception to the described procedure was the selection of one school each from cells D, E, and L, even though the proportionate number could not be rounded logically to one. The investigator believed that these three cells should be represented in the final sample.

The sum of the schools to be selected from the cells was 101. It was decided to use this number in the final sample.

A table of random numbers (4, *pp. 262-264*) was applied to each cell until the proper number of cards had been selected. The randomly selected cards in each cell were turned face up. Identifying numbers on the cards were matched with the number given each school earlier. The 101 sample schools were thus identified.

ANALYSIS OF THE SAMPLE

The selected sample is representative with respect to size of school and size of community. An analysis of the sample with respect to other independent variables was made.

In each analysis, the χ^2 test was employed to test the hypothesis that the table of frequencies for the classification in the selected sample was that which might generally be expected in a random sample selected from the 1,040 schools in Ohio.

Adopting the 5-per-cent level of confidence for rejection, the hypotheses were retained for geographic area, accreditation, and for the type of school

TABLE 1

Stratification of Ohio Schools by School and Community Size with Number of Schools To Be Selected for Sample, 1952

School size	Community Size					Total schools
	0-499	500-999	1,000-4,999	5,000-19,999	20,000 and up	
0-199	*A* N = 237 9.6% = 22.36 22 schools	*B* N = 169 9.6% = 16.22 16 schools	*C* N = 117 9.6% = 11.23 11 schools	*D* N = 5 9.6% = .48 1 school	*E* N = 2 9.6% = .19 1 school	51
200-499	*F* N = 24 9.6% = 2.30 2 schools	*G* N = 68 9.6% = 6.53 7 schools	*H* N = 183 9.6% = 17.57 18 schools	*I* N = 46 9.6% = 4.42 4 schools	*J* N = 9 9.6% = .86 1 school	32
500-999	*K* N = 1	*L* → N = 2+1 9.6% = .29 1 school	*M* N = 18 9.6% = 1.73 2 schools	*N* N = 56 9.6% = 5.38 5 schools	*O* N = 34 9.6% = 3.26 3 schools	11
1,000 and up	*P* No schools	*Q* No schools	*R* N = 1	*S* → N = 7+1 9.6% = .77 1 school	*T* N = 65 9.6% = 6.24 6 schools	7
Total schools	24	24	31	11	11	101

district[3] from each geographic area. The hypothesis was rejected for the type of school district on a state-wide basis.

Findings

Space will not permit a detailed discussion of all the findings in this study. The reader is referred to the original dissertation by the writer for additional information (6).

ITEM ANALYSIS

The *Health and Physical Education Score Card No. II* (3) is composed of ten areas. Each area contains ten items which are to be scored from 0 to 3, depending upon how well a school compares with the standards established for each item.

Rank Order. By placing the 100 items in rank order by mean item score, it was found that no item received a perfect score in all the 101 Ohio schools surveyed. Eight items had a mean score above 2, while 36 items had a mean score above 1. There were seven items which received a score of 0 in each of the 101 schools evaluated.

All items from Area V, Swimming Pool, and Area VIII, Modified-Individual (Corrective) Activities, were among the lowest 25 mean item scores.

In general, the mean scores for the highest 25 items in each school-size category appear to increase as school size increases up to an enrollment of 1,000. There seems to be a slight decline in some mean item scores for schools with an enrollment above 1,000.

Discriminative Ability. The Votaw Curve technique (11) was used to determine the ability of each item to discriminate between high and low area scores and high and low total scores.

There were 52 items which discriminated between high and low area scores and 44 items which discriminated between high and low total scores. In general, the most highly discriminative item in each area also displayed discriminative ability when plotted by total score.

Recommended Ohio Standards. Few of the health and physical education standards recommended by the Ohio State Department of Education (2) are being met by the schools. All schools sponsoring a program of interscholastic athletics appear to be following the Rules and Bylaws of the Ohio High School Athletic Association.

SCORE CARD DIVISION ANALYSIS

The ten score-card areas constitute one method of score-card division. These ten areas are: I. Program of Activities; II. Outdoor Areas; III. Indoor Areas; IV. Locker and Shower Areas; V. Swimming Pool; VI. Supplies and Equipment; VII. Medical Examinations and Health Service; VIII. Modified-

[3]Ohio has three types of school districts: county; exempted village (those communities with a population between 3,000-5,000 who choose to withdraw from the county district); and city (those communities with a population over 5,000). This study divided county schools into a consolidated group and a non-consolidated group.

Individual (Corrective) Activities; IX. Organization and Administration of Class Programs; and X. Administration of Intramural and Interschool Athletics.

The title of the score card implies that a comparison might be made between the health program and the physical education program.

It also would appear that the score card could be divided into two parts: (a) facilities, and (b) program content or administrative procedures.

Area Analysis. Area II, Outdoor Areas, had the highest mean area score with 11.8 of a possible 30 points. Following in order behind Area II were: Area IX (11.7); Area X (10.8); Area VI (10.6); Area IV (10.2); Area III (9.2); Area VII (8.6); Area I (7.1); Area V (.6); and Area VIII (.3).

The Pearson product-moment correlation coefficient was used to determine a measure of the relationship between total scores, selected independent variables, and area scores. Biserial correlation was used for the two dichotomous variables, accreditation and consolidation.[4] The correlation matrix is shown in Table 2.

TABLE 2

Correlations Between Area Scores, Total Score, and Selected Variables

Variable	Area							
	I	II	III	IV	VI	VII	IX	X
Total Score	.85	.40	.88	.77	.61	.64	.69	.59
Accreditation	.56	.25	.67	.58	.52	.50	.49	.44
School Size	.58	.11	.58	.50	.48	.41	.44	.41
Town Size	.33	.13	.40	.27	.17	.19	.28	.15
Consolidation	.10	−.14	−.01	.29	−.10	−.07	.24	−.01

Table 2 reveals that Area III, Indoor Areas, with an r of .88, correlated highest with total score. Area I, Program of Activities, and Area IV, Locker and Shower Areas, followed with r's of .85 and .77 respectively. This is interpreted to mean that, in general, schools more nearly meeting acceptable standards for indoor areas, program of activities, and locker and shower areas tend to have a higher quality health and physical education program.

The relationship between accreditation and area scores is generally higher than for school size, town size, or consolidation. Since accredited schools are generally required to meet certain standards in health and physical education prior to receiving an accredited rating, it would appear logical to expect some positive relationship to exist.

The r found between school size and the several area scores is interpreted to mean that, with the exception of Area II, Outdoor Areas, there is a definite tendency for area scores to increase as school size increases. This is similar to the findings between school size and item scores noted in the item analysis section.

[4] r_{bis} between consolidated and non-consolidated schools.

In general, there would appear to be no systematic increase in area scores when town size increases.

All correlations between consolidation and area scores are insignificant. They are slightly negative in five of the eight areas. This is interpreted to mean that, in general, there is no appreciable difference between area scores of county consolidated and county non-consolidated schools.

The Health Program and the Physical Education Program. A committee of five members of the School of Health, Physical Education, and Recreation faculty of Indiana University selected 39 of the 100 score-card items as being more closely associated with health and safety than with physical education (agreement between at least three of the five was determined necessary for retention as a health and safety item). Seven additional items were determined by the committee to be equally concerned with health and safety and physical education.

The statistic χ^2 was employed to test the hypothesis that the percentage of the total Ohio sample points actually attained by each program was not significantly different from the percentage of points which would have been proportionate.

It was found that the value of χ^2 for this test (1.020, 1 d.f.) would be exceeded in more than 30 per cent of the time by chance alone. Thus, it appears that there is no significant difference in the attainment of the two programs by the 101 selected Ohio schools.

Facilities and Program Content or Administrative Procedures. A comparison between facilities and program content or administrative procedures was made in a manner similar to that described in the previous section.

It was found that the value of χ^2 for this test (.401, 1 d.f.) would be exceeded in more than 50 per cent of the time by chance alone. Thus, it appears that there is no significant difference in the attainment between these two score-card divisions.

TOTAL SCORE ANALYSIS

The total scores ranged from 31 to 154 of the possible 300 points. Other statistical characteristics of the scores were: $M = 80.83$; $Mdn = 77$; $\sigma = 22.06$; and $\sigma_M = 2.20$.

The cumulative frequency curve in Figure I graphically represents the distribution of total scores for the 101 sample schools. Ohio schools wishing to survey themselves may compare their status with these norms.

Analysis of Group Means. The analysis of variance technique described by Lindquist (4, *pp. 87-179*) was used to test the hypothesis that the differences in means of the groups within each classification of the variables analyzed might be attributed to fluctuations in random sampling and that there were no real differences.

Bartlett's formula (4, *pp. 99-100*) was used to test the hypothesis of homogeneous variance in all analyses using the analysis of variance technique.

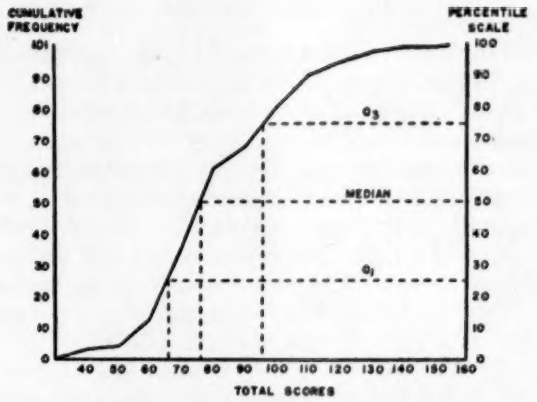

FIGURE I. Cumulative Frequency Curve Showing Percentile Ranks Corresponding to Total Health and Physical Education Score Card No. II Scores Made by 101 Selected Ohio Secondary Schools, 1952-53.

Using the 5-per-cent-level of confidence for rejection, the hypothesis was retained for all analyses except for accredited and non-accredited schools.

The ratio (F) between the category variance and the within group variance was significant at better than the 1-per-cent-level of confidence in each analysis.

Table 3 shows the data used in the school-size analysis.

TABLE 3

Analysis of Variance of Total Scores by School Size

Category	N	ΣX	Mean	ΣX^2
0 - 199	51	3,548	69.806	259,262
200 - 499	32	2,598	81.188	217,938
500 - 999	11	1,261	114.636	149,203
1,000 - up	7	757	108.143	82,663

For the school-size categories, subsequent t tests revealed significant differences, at better than the 1-per-cent-level of confidence, between the means of each category except the 500-999 and the 1,000-and-up categories. The latter difference was not significant at the 5-per-cent-level of confidence. These results are similar to those found when school size was used as a factor for both score-card items and areas.

Table 4 gives the data used in the type of school-district analysis.

TABLE 4

Analysis of Variance of Total Scores by Type of School District

Category	N	ΣX	Mean	ΣX^2
County	58	4,201	72.431	322,441
County (consolidated)	13	970	74.615	73,916
Exempted Village	13	1,161	89.308	111,093
City	17	1,832	107.765	201,616

The t test for the data in Table 4 revealed no significant difference between the means of the county consolidated and county non-consolidated schools. Both the exempted village and city schools had significantly higher means than either category of county schools.

The difference in each case was significant at better than the 1-per-cent-level of confidence except for the difference between the exempted village and consolidated schools. In this instance, the level of confidence was 5 per cent.

The t test revealed a significant difference at better than the 1-per-cent-level of confidence between the city and exempted village schools.

Table 5 indicates the data used in the geographic area analysis.

TABLE 5

Analysis of Variance of Total Scores by Geographic Area

Category	N	ΣX	Mean	ΣX²
Southeast	20	1,258	62.900	85,724
Northwest	24	1,905	79.375	156,845
Southwest	30	2,632	87.733	247,624
Northeast	27	2,369	87.741	218,873

Subsequent t tests between means in Table 5 revealed significant differences, at better than the 1-per-cent-level of confidence, between the southeast geographic area and each other area. No other significant differences were found.

Table 6 shows the data used in the accreditation analysis.

TABLE 6

Analysis of Variance of Total Scores by Accreditation

Category	N	ΣX	Mean	ΣX²
Non-accredited	53	3,644	68.755	262,264
Accredited	48	4,520	94.167	446,802

Since the data in Table 6 were found to lack homogeneous variance, Snedecor's technique was used for testing the difference between means with no hypothesis about the variance (9). This test revealed a significant difference between the two means at better than the .1-per-cent level of confidence.

Influence of Selected Variables Upon Total Score. In an attempt to determine the relative influence of selected variables upon total score, three different methods were used. They were: causal analysis; the ratio between the "low" mean and the "high" mean of each variable in the causal analysis technique(1); and correlation.

Accreditation was found to have the most influence upon total score by all three methods. The rank order of the other variables by degree of influence was: school size; geographic area; town size; and consolidation.

Summary and Conclusions

A controlled sample of 101 Ohio public secondary schools was evaluated by means of the *Health and Physical Education Score Card No. II* (3). The selected sample was representative with respect to size of school and size of community. It was also revealed that the table of frequencies for geographic area, accreditation, and type of school district for each geographic area in the selected sample was that which might generally be expected in a random sample selected from the 1,040 schools in Ohio.

The following conclusions seem justified on the basis of the results:

Schools with an enrollment of 500-999 students have a significantly better health and physical education program than either the 0-199 or 200-499 school-size categories. Enrollment above 1,000 students tends to have a slight negative influence upon item scores, area scores, and total score.

Accreditation and school size have considerably more influence upon the quality of the health and physical education program than geographic area, town size, or consolidation.

The present consolidation movement among the county schools of Ohio has not resulted in a significant improvement over the county non-consolidated schools with respect to health and physical education programs.

The 101 selected Ohio secondary schools more closely follow the mandatory Rules and Bylaws of the Ohio High School Athletic Association than the recommended standards of the Ohio State Department of Education.

REFERENCES

1. Fritz, Harry G. *An Evaluation of the Boys' Health and Physical Education Program in Selected White Secondary Schools of Missouri.* Doctor's dissertation, Indiana Universilty, Bloomington, 1954. p. 119.
2. Landis, Paul E., and P. C. Bechtel. *Ohio High School Standards for Health, Physical Education, Recreation, and Safety.* Columbus: Ohio State Department of Education, 1949.
3. LaPorte, William R. *Health and Physical Education Score Card No. II.* Los Angeles: University of Southern California Press, 1947.
4. Lindquist, Everet F. *Statistical Analysis in Educational Research,* Chicago: Houghton Mifflin Co., 194
5. Ohio State Department of Education. *Educational Directory.* Columbus: Ohio State Department of Education, 1951.
6. Owens, Laurence E. *An Analysis of the Health and Physical Education Program for Boys in Selected Ohio Secondary Schools.* Doctor's dissertation, Indiana University, Bloomington, 1955.
7. Rand McNally and Co. *Rand McNally Commercial Atlas and Marketing Guide.* 82nd edition. New York: Rand McNally and Co., 1951.
8. Research Division, National Education Association. Teaching Load in 1950. *Research Bulletin,* **29**: 1-51, 1951.
9. Snedecor, George W. *Statistical Methods.* 4th ed. Ames: Iowa State College Press, 1946. p. 84.
10. U. S. Bureau of the Census. *United States Census of Population: 1950.* Washington, D. C.: U. S. Govt. Printing Office, 1951.
11. Votaw, David F. Graphical Determination of Probable Error in Validation of Test Items. *Journal of Educational Psychology,* **24**: 682-686, 1933.

(Submitted 1/10/55)

Comparisons of Reaction-Time Measures to a Visual Stimulus and Arm Movement

A. T. SLATER-HAMMEL
Indiana University
Bloomington, Indiana

Abstract

The purposes of this study were (1) to compare reaction-time measures for arm displacement and a visual stimulus, and (2) to compare reaction-time measures for selected groups of varsity athletes, physical education majors, music majors, and liberal arts majors. Analysis of the data revealed that only a modest relationship existed between the two reaction-time measures. Significant differences in reaction time were found among the several groups for both reaction-time measures.

THE FACT THAT man's kinesthetic sense is an important determinant in the control of body movements has been long recognized and has led to numerous investigations of this sense. In the main, the bulk of research in this area has been and continues to be concerned with discriminations of muscular force, positioning of body parts, and direction of movement. A few recent investigations, however, have gotten off the beaten path and have attempted to obtain estimates of kinesthetic reaction time.

Vince and Hicks (11, 5), for example, determined the time required for subjects to correct movement of a lever when sudden changes in resistance were imposed upon the lever system. These investigators found that the corrective movements required from 0.16 sec. to 0.30 sec. Chernikoff and Taylor (1) found that reactions to sudden movement of the arm and sudden stopping of an arm movement required approximately 0.15 sec. and 0.12 sec. respectively. Data have been presented by Denenberg (2) indicating that reactions to sudden arm drop were delayed by approximately 0.16 sec.

Naturally, there has been no claim that these estimates of kinesthetic reaction time are pure measures. The investigators cited above have been the first to recognize that their experimental procedures involved more than kinesthetic stimulation. In producing sudden changes in the resistance of a lever system or sudden changes in the momentum of a subject's arm, tactual stimuli are unavoidably introduced. Under such conditions, it is impossible to know if the subject is responding to kinesthetic stimulation, to tactual

stimulation, or to a complex of kinesthetic and tactual stimuli.[1] But in spite of this limitation, an interesting and a possibly significant form of human response has become available for investigation.

The study reported here is concerned primarily with one form of this new response—the speed with which subjects can react to the sudden displacement of an arm. Specifically, the purposes of this investigation were (1) to compare reaction times for arm displacement and a visual stimulus, and (2) to compare reaction times for selected groups of subjects. Since the experimental procedures for obtaining measures of the speed with which subjects could react to the sudden displacement of an arm involved both tactual and kinesthetic stimuli, the measure will not be called "kinesthetic reaction time," as has been done by earlier investigators; the measures simply will be referred to as "reaction time to arm movement."

Method

Apparatus for Reaction Time to Arm Movement. The apparatus for the determination of reaction time to arm movement was similar to that described by Denenberg (2). Briefly, it consisted of an electromagnet with an energizing source of direct current, an experimenter's switch, a soft iron bar riveted to a leather wristband, a telegraph key, and a chronoscope. The electromagnet was mounted upon a horizontal framework so that it was suspended 40 inches above the floor. The telegraph key was clamped to the edge of a table which was positioned to the right of the electromagnet. Since the chronoscope was not absolutely noiseless, this unit and the experimenter's switch were located in an adjacent room.

To obtain measures of reaction time, the iron bar was attached to the subject's wrist by means of the wristband. The subject then sat on a stool beside the table, rested his right arm on the table top, and depressed the telegraph key with his fingers. From this position, the subject abducted his left arm and brought the iron bar in contact with the magnet. Once contact was made with the magnet, the subject's arm was held fast in the abducted position by magnetic attraction. With the subject in this starting position, the operation of the apparatus was such that the following actions occurred: (1) opening of the experimenter's switch rendered the magnet inoperative and caused the subject's left arm to fall, (2) movement of the iron bar from the

[1]Although it appears to have been unrecognized, tactual stimulation of some sort is generally involved in the more traditional tests of kinesthesis. When, for example, the subject is required to exert force against a scale, contact of the body part with the scale introduces a tactual stimulation. In tests where the blindfolded subject is required to move a body part to a particular point in space, the test trial is usually preceded by a conditioning trial in which the experimenter moves the subject's body part, and this introduces tactual stimulation. Precisely what effect the differences in stimulation between conditioning and test trials has upon the subject's performance is unknown; however, since the stimulating conditions between conditioning and test trials are different, there is little reason to expect the subject's responses to be exactly alike.

magnet poles started the chronoscope, and (3) movement of the subject's fingers off the telegraph key stopped the chronoscope.

Apparatus for Reaction Time to a Light Stimulus. The apparatus for measurement of reaction time to a visual stimulus consisted of a 105-115 volt neon glow lamp, a telegraph key, an experimenter's switch, and a chronoscope. The lamp was secured to one end of a table so that it rested 15 inches above the table top. The telegraph key was located at the other end of the table. This key was not secured to the table and could be moved in any position found to be most comfortable for the subject. In order to eliminate undesired sounds, the chronoscope and the experimenter's switch were located in an adjacent room.

The several components were wired together so that closing of the experimenter's switch simultaneously activated the lamp and chronoscope. Movement of the subject's fingers off the telegraph key stopped the chronoscope.

Procedures. A standardized explanation and demonstration was given before each reaction-time test. Following this presentation, the subject was permitted to practice the desired reaction.

In the reaction-time test to arm movement, the subject was first given practice in relaxing his suspended arm and allowing it to drop freely when released by the electromagnet. No attempt was made to have the subject react to the movement of his arm during these preliminary trials. When it became evident that the subject could relax his arm, he was blindfolded and given five practice reaction-time trials. During these trials, the subject was instructed to concentrate upon the movement of his arm and to react as soon as he became aware of any movement. At the completion of these trials, the regular test was started.

In the test of reaction time to a light stimulus, the subject was permitted to place the telegraph key in any position he felt most comfortable for operation. Following this adjustment, the subject was given five practice trials before undertaking the regular test trials. Regardless of hand preference, the subject was required to operate the telegraph key with the fingers of his right hand.

In both reaction-time tests, a verbal preparatory signal preceded presentation of the response stimulus. The time interval between the verbal signal, "Ready," and the response stimulus was systematically varied during a series of trials. The intervals were approximately one, two, and three seconds, arranged in random order.

Since the experimenter's switch and the chronoscope used for both tests were located in an adjacent room, an assistant was employed to operate the controls and tabulate reaction-time measures. A speaker and microphone system was used for communication between the two rooms.

Each subject completed a total of 50 trials in each reaction-time test. The order in which the two tests were completed was systematically varied for all groups of subjects. A two-minute rest interval was given after the completion of 25 trials, and an interval of approximately five minutes intervened between the two tests.

Subjects. Eighty male university students ranging in age from 18 to 27 years served as subjects. These students were obtained on a voluntary basis and consisted of the following undergraduate groups: (I) 20 varsity athletes, (II) 20 physical education majors who were not varsity athletes, (III) 20 music majors who were not varsity athletes, and (IV) 20 liberal arts majors who were not varsity athletes. These particular groups of students were selected because previous studies have suggested that athletes, physical education majors, and musicians tend to be superior to other students in tests of kinesthetic perception (4, 6, 8, 9, 12).

Analysis of Data. For the experimental design involving repeated measurements of independent groups of subjects, an analysis of variance was employed to provide an over-all evaluation of the data. The analysis was carried out according to procedures outlined by Lindquist (7; *pp. 267-273*) for the Type I design. A description and a numerical example of this analysis also has been provided by Edwards (3; *pp 288-297*).

When justified by the F values obtained in the analysis of variance, t tests were employed to evaluate the differences between means for both the main and simple effects of the experimental variables. The appropriate error terms for these tests have been summarized by Lindquist (7; *pp. 270-273*).

The product-moment method of correlation was used to obtain an estimate of the relationship between the two reaction-time measures.

Prior to collection of the data, it was decided to require a significance level of 1 per cent for rejection of the null hypothesis for all F and t values. Having accepted this class of results, no attempt was made to follow the comon practice of indicating the maximum level of significance for each F and t value.[2]

Results

Main Effects and Interaction. The variance analysis of repeated measures on independent groups of subject revealed F values of 765.63 for between reaction-time measures, 7.54 for the interaction of reaction-time × groups, and 6.96 for between groups of subjects. These F values were all significant

[2]An excellent discussion of the justification for these procedures has been presented by Lindquist (7, *pp. 90-95*). As indicated by Lindquist, the practice of indicating individual probabilities for each comparison is objectionable because (1) acceptance or rejection of the null hypothesis is categorical and not a matter of degrees, (2) it leads to constant shifts in risks of a Type II error, and (3) it frequently results in attempts to use individual probabilities as a means of estimating the importance of differences between means.

beyond the 1% level, and it was concluded that the null hypothesis for the two main effects and the interaction of these effects was not tenable.

For all subjects, the mean reaction-time to arm movement was 148 ms., to a visual stimulus 197 ms. Since the F value for the variance attributable to between reaction-time measures resulted in rejection of the null hypothesis, it was concluded that reaction-time to arm movement was significantly shorter than reaction-time to a visual stimulus.

The significant F value for interaction was interpreted as indicating a differential in response of subject-groups to the two reaction-time situations. Some insight into this differential in response will be gained during presentation of the simple effects involving comparisons of reaction-time measures for groups of subjects.

The sgnificant F value for between groups of subjects was interpreted as indicating that at least some of the differences in over-all reaction time between the several groups were significant. Comparisons of the differences between groups are summarized in Table 1. From the t values obtained, it would appear that the group fell into two main sets, with varsity athletes constituting one set and music majors and liberal arts majors the other. The mean reaction time for varsity athletes was significantly shorter than that for music majors and liberal arts majors, but the difference in mean reaction time for music majors and liberal arts majors was not significant. The physical education majors fell between these two sets, and their position was ambiguous.

TABLE 1

Comparisons of Over-All Reaction Time for Selected Groups of Subjects

Group	Mean Reaction Time (ms.)	Diff.	S.E. of Diff.	t	P
Varsity Athletes	160				
Phys. Educ. Majors	170	10	5.38	1.86	Below 1%
Varsity Athletes	160				
Music Majors	175	15	5.38	2.79	Above 1%
Varsity Athletes	160				
Lib. Arts Majors	184	24	5.38	4.46	Above 1%
Phys. Educ. Majors	170				
Music Majors	175	5	5.38	—	———
Phys. Educ. Majors	170				
Lib. Arts Majors	184	14	5.38	2.60	Below 1%
Music Majors	175				
Lib. Arts Majors	184	9	5.38	1.67	Below 1%

In an effort to resolve the ambiguous position of physical education majors, the data were re-examined according to Tukey's (10) procedure for comparing individual means. The results of this analysis may be summarized as follows:

1. The gap test failed to separate any of the groups.
2. The straggler test separated varsity athletes from physical education majors, music majors, and liberal arts majors.
3. The straggler test failed to separate any group among physical education majors, music majors, and liberal arts majors.
4. The excessive variability test for physical education majors, music majors, and liberal arts majors was not significant.

On the basis of these findings, it does not seem unreasonable to place physical education majors in the set including music majors and liberal arts majors. The mean reaction times for the members of this set can be considered alike among themselves, but significantly longer than the mean reaction time for varsity athletes.

Simple Effects. The simple comparisons of reaction-time measures to a visual stimulus and to arm movement are summarized in Table 2. As shown in Table 2, reaction time to arm movement for the several groups ranged from 141 ms. to 160 ms., and reaction time to a light stimulus ranged from 179 ms. to 209 ms. Since all t values were significant beyond the 1 per cent level, it is evident that reaction time to arm movement was significantly shorter for all groups of subjects than reaction time to a visual stimulus.

TABLE 2

Comparisons of Reaction Time Measures to a Visual Stimulus and Arm Movement for Selected Groups of Subjects

Group	Visual Stimulus (ms.)	Arm Movement (ms.)	Diff.	S.E. of Diff.	t	P
Varsity Athletes	179	141	38	3.48	10.92	Above 1%
Phys. Educ. Majors	196	145	51	3.48	14.65	Above 1%
Music Majors	205	145	60	3.48	17.24	Above 1%
Lib. Arts Majors	209	160	49	3.48	14.08	Above 1%

The comparisons of reaction-time measures to arm movement between the several groups is presented in Table 3. It is evident from an inspection of Table 3 that the groups fell into two clear-cut sets, with varsity athletes, physical education majors, and music majors comprising one set and liberal arts majors the other set. Although the differences in reaction time among members of the first set were not significant, they were all significantly faster than liberal arts majors.

Table 4 summarizes the comparisons of reaction-time measures to a visual stimulus for the several groups of subjects. From the t values obtained, it is

TABLE 3

Comparisons of Reaction Time Measures to Arm Movement for Selected Groups of Subjects

Group	Reaction Time (ms.)	Diff.	S.E. of Diff.	t	P
Varsity Athletes	141				
Phys. Educ. Majors	145	4	5.46	—	——
Varsity Athletes	141				
Music Majors	145	4	5.46	—	——
Varsity Athletes	141				
Lib. Arts Majors	160	19	5.46	3.48	Above 1%
Phys. Educ. Majors	145				
Music Majors	145	0	5.46	—	——
Phys. Educ. Majors	145				
Lib. Arts Majors	160	15	5.46	2.75	Above 1%
Music Majors	145				
Lib. Arts Majors	160	15	5.46	2.75	Above 1%

TABLE 4

Comparisons of Reaction Time Measures to a Visual Stimulus for Selected Groups of Subjects

Group	Reaction Time (ms.)	Diff.	S.E. of Diff.	t	P
Varsity Athletes	179				
Phys. Educ. Majors	196	17	6.28	2.71	Above 1%
Varsity Athletes	179				
Music Majors	205	26	6.28	4.14	Above 1%
Varsity Athletes	179				
Lib. Arts Majors	209	30	6.28	4.78	Above 1%
Phys. Educ. Majors	196				
Music Majors	205	9	6.28	1.43	Below 1%
Phys. Educ. Majors	196				
Lib. Arts Majors	209	13	6.28	2.07	Below 1%
Music Majors	205				
Lib. Arts Majors	209	4	6.28	—	——

evident that the groups could again be classified into two clear-cut sets. The reaction time for varsity athletes, constituting one set, was significantly shorter than the set including physical education majors, music majors, and liberal arts majors; but the differences among members of the latter set were not significant.

It was noted in an earlier comment that the significant F value for interaction indicated a differential in response of groups to the two reaction time situations. Some understanding of this differential in response may be gained through consideration of the performance of physical education majors and music majors in the two reaction-time situations. In the comparisons of reaction time to arm movement, physical education majors and music majors were members of the set constituting fast reactors. But in comparisons of reaction time to a visual stimulus, these two groups were members of the set constituting slow reactors. In other words, the position of physical education majors and music majors depended upon the particular reaction-time situation.

Relationship Between Reaction Time Measures. The correlation between the two reaction time measures was found to be 0.562. Although this value was significant beyond the 1 per cent level, it can be said to indicate no more than a modest relationship between reaction time measures.

Discussion

In the relatively new reaction-time situation involving response to arm movement, it is of some interest to note that the estimates found in this investigation ranged from 141 ms. to 160 ms. This range appears to be in general agreement with data reported by Chernikoff and Taylor (1) and Denenberg (2). These investigators reported reaction times ranging from 149.6 ms. to 158.34 ms.

The decisive findings in favor of reaction time to arm movement over reaction time to a visual stimulus leave little doubt that the former is significantly shorter than the latter. This, of course, was not unexpected. From the numerous published estimates of reaction time to a visual stimulus, and from the available estimates of reaction time to arm movement, it would have been most surprising to obtain other results.

Although the results of this investigation revealed significant differences among the several groups of subjects in both reaction-time situations, there is considerable uncertainty as to the extent one may generalize on the basis of these findings. For one thing, it is operationally difficult to justify any sweeping assumption of representativeness for the groups involved. This investigator, for instance, would not care to subscribe to the assumption that the collection of varsity athletes selected from Indiana University are represenative of all varsity athletes in reaction time, or in any other characteristic. Yet, unless an investigator can claim representativeness for his samples, generalization is an extremely hazardous procedure.

A second and equally powerful deterrent to generalization lies in the possibility that the groups of subjects involved might not even be representative of the classes to which they belong at Indiana University. It was not possible to choose the subjects in each group at random. University students are busy people, they have a variety of interests to occupy their time, and they frequently are not interested in serving as subjects. As is the situation in most investigations on human subjects, it was necessary to use only those students who were willing to co-operate and come to the laboratory.[3] To what extent this method of selection biased or distorted the experimental results is unknown. And the uncertainty on this point makes one hesitate to take the inductive leap, to pronounce the careless generalization.

Because of the limitations in sampling, which are by no means unique to this investigation, it is recommended that group differences in reaction time simply be accepted as suggestive of real differences.

Conclusions

An experiment was conducted to determine the reaction time to a visual stimulus and to arm movement for selected groups of subjects. The groups included varsity athletes, physical education majors, music majors, and liberal arts amjors. For the subjects used in this investigation, it was concluded that:

1. Reaction time to arm movement was significantly shorter than reaction time to a visual stimulus for all subjects and for all groups of subjects.
2. Varsity athletes had a significantly shorter over-all reaction time than physical education majors, music majors, and liberal arts majors. Other group differences were not significant.
3. Varsity athletes, physical education majors, and music majors had a significantly shorter reaction time to arm movement than liberal arts majors. Other group differences were not significant.
4. Varsity athletes had a significantly shorter reaction time to a light stimulus than physical education majors, music majors, and liberal arts majors. Other group differences were not significant.
5. Only a modest relationship existed between the two reaction-time measures.

REFERENCES

1. Chernikoff, R. and F. V. Taylor. Reaction Time to Kinesthetic Stimulation Resulting from Sudden Arm Displacement. *Journal of Experimental Psychology*, **43**: 1-8, 1952.
2. Denenberg, V. H. A Simplified Method of Measuring Kinesthetic Reaction Time. *American Journal of Psychology*, **66**: 309-311, 1953.
3. Edwards, A. L. *Experimental Design in Psychological Research*. New York: Rinehart and Co., Inc., 1950, p. 446.

[3]An example of the common difficulty in securing human subjects might be illustrated by noting that it required interviews with approximately 60 music majors to obtain a pool of 20 subjects.

4. FILER, P. W. *The Relationship Between Kinesthetic Judgment and Success as an Instrumental Musician.* Unpublished Master's thesis, Pennsylvania State College, State College, 1934.
5. HICKS, W. E. Reaction Time for the Amendment of a Response. *Quarterly Journal of Experimental Psychology*, **1**: 175-179, 1949.
6. KERR, W. H., and J. D. WEINLAND. Muscular Perceptivity as a Trade Test. *Journal of Applied Psychology*, **17**: 550-558, 1933.
7. LINDQUIST, E. F. *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin Co., 1953.
8. STEVENS, M. *The Measurement of Kinesthesis in College Women.* Unpublished Doctor's dissertation, Indiana University, Bloomington, 1950.
9. TAYLOR, W. J. *The Relation Between Kinesthetic Judgment and Success in Basketball.* Unpublished Master's thesis. Pennsylvania State College, State College, 1933.
10. TUKEY, J. W. Comparing Individual Means in the Analysis of Variance. *Biometrics*, **5**: 99-114, 1949.
11. VINCE, M. A. Corrective Movements in a Pursuit Task. *Quarterly Journal of Experimental Psychology*, **1**: 85-103, 1948.
12. WEIBE, V. R. *A Study of Tests of Kinesthesis.* Unpublished Master's thesis, State University of Iowa, Iowa City, 1951.

(Submitted 1/28/55)

Notes and Comments

NOTES

Russian Studies of Vision in Relation to Physical Activity and Sports

ASHTON GRAYBIEL, M.D.
U. S. Naval School of Aviation Medicine
Pensacola, Florida
ERNST JOKL, M.D., AND CLAUDE TRAPP, M.D.
University of Kentucky
Lexington, Kentucky

Part I. The Visual Apparatus

AT THE OCCASION of the 1952 Olympic Games in Helsinki, the Finnish Government, in collaboration with the International Olympic Committee, the Finnish Association of Sports Medicine and the Federation Internationale Medicine Sportive, organized a symposium to discuss problems of normal and clinical physiology of athletics. Dr. M. J. Karvonen[1] of Helsinki was Organizing Secretary. Prior to the Olympic Games high level agreement had been reached by the governments of the participating countries and it was therefore possible for physicians and physiologists from both sides of the Iron Curtain to meet. Dr. Jokl presided at the sectional meeting of the Symposium in which the Russian delegation under Prof. Kukolewsky participated. The latter presented the Chair with a copy of a new book by A. N. Krestovnikov entitled *Studies in Physiology of Exercise* which had been published by the U.S.S.R. Government Printing Offices in Moscow.

Since the publication in 1938 of Percy M. Dawson's *Soviet Samples. Diary of an American Physiologist* (Edward Brothers, Inc., Ann Arbor, Mich.), this was the first opportunity to become acquainted with work in a field of study which is of increasing significance. Scope and quality of Dr. Krestovnikov's book appeared at cursory examination to be so interesting that it was judged desirable to have its contents available in English.

The Director of Naval Intelligence, Office of the Chief of Naval Operation, Department of the Navy, Washington, D. C., kindly concurred with a request that the book be translated. The necessary arrangements were completed in the summer of 1955. The following article reviews the first chapter. It is concerned with the mutual relationship between a number of physicological components of vision and of exercise and training.

Visual Acuity

Two investigations are concerned with the effect of exercise on visual acuity. Tests were conducted before and after a 1000-m. race. In 27% of the athletes visual acuity remained unchanged after the race while in 73% visual acuity increased "by as much as 45%." In ten runners, acuity was checked every ten minutes following the exercise. Pre-race acuity values were reestablished in two cases within 50 minutes after the race; in four cases in one and one-half hours; and in another four cases within two hours. Greatest sharpness of vision was found either immediately after the race or ten minutes later.

The author remarks that these findings require amplification and corroboration. He cites additional evidence which suggests that differences in the running distance have a bearing upon the results.

Accommodation

Accommodation, i.e., the ability of the eye to adjust its focus to various distances, was measured on 100 well-trained athletes. The usual method of testing is to bring small print toward the eye (after refractory disturbances are corrected through glasses) until a point is reached when the print blurs. No deviations from normal standards were found in the athletes.

---

[1]M. J. Karvonen, *Sport Medicine, Proceedings of the International Symposium of the Medicine and Physiology of Sports and Athletics at Helsinki*, 17-18. July 1952. Publisher: Finnish Association of Sports Medicine.

Visual Fields

In another group of athletes visual fields were plotted. Visual fields measure the extent of peripheral sight. The eyes are fixed on a point in the central field of vision while a test target, usually a small bead, approaches the center from outside the range of vision. The first glimpses of the target define the border line of the visual field. Normally, visual fields are larger for white than for any color. Blue is next largest, followed by red. The smallest visual fields are measured for green.

Visual fields of 70 basketball players, 162 soccer players, 14 hockey players, and 20 tennis players were plotted. There were no differnces from the established norm with white, red, and blue test objects. However, visual fields for green were considerably larger in the athletes than they were in the controls:

In a separate study, visual fields were measured after athletes competition. In the majority of cases, a post-exercise increase was found in the size of the visual fields, most markedly so for green, less so for the other colors and for white.

Ocular Muscle Balance

Ocular muscle balance studies were conducted with 401 individuals. These studies consisted of heterophoria measurements, performed at both 25 cm. ("near" vision) and 5 m. ("far" vision). Normally, both eyes are directed at the same object with the result that the two images are "fused" into a single visual picture. Heterophoria measurements reveal the axis positions of the two eyes when each eye is confronted with a separate image, or when one eye is covered. Such eyes are said to be in the "fusion-free" position or in the "position of rest." The deviation between the directions of the eyes in the "fusion-free" position represents heterophoria, which is measured in prism diopeters. One prism diopter is almost exactly equal to one-half degree. Exact zero heterophoria (orthophoria) is uncommon. Normal values of heterophoria at 5 m. vary from about three prism diopters esophoria (tendency to converge the directions of the eyes) to about three prism diopters exophoria (tendency to diverge). At 25 cm. normal values range from about one prism diopter esophoria to seven prism diopters exophoria.

The phoria measurements of a representative group of 194 untrained individuals were compared with those of 25 tennis champions (Table 1). *The champion athletes had significantly lower degrees of heterophoria or, to use different terminology, were more orthoporic than the non-athletes.*

TABLE 1

Comparison of Phoria Measures

At 5 m. distance

	orthophoria per cent	percent heterophoria: 0.1-1Δ	1.1-3.0Δ	3.1-5Δ	5.1-8Δ	8.1-11Δ	11,1-14Δ
194 untrained students	40	40.1	13.0	4.2	1.5	1.0	....
25 tennis champions	72	20.0	8.0	....	....	....	....

At 25 cm. distance

	orthophoria per cent	percent heterophoria: 0.1-1Δ	1.1-3.0Δ	3.1-5Δ	5.1-8Δ	8.1-11Δ	11,1-14Δ
194 untrained students	1.5	5.9	26.4	26.2	23.2	11.3	4.6
25 tennis champions	8.0	56.0	20.0	8.0	8.0	....	....

Other tests showed that trained subjects tend to be more orthophoric than untrained ones. Of a trained group, 68% were essentially orthophoric, of an untrained 40%. The former comprised over 500 athletes, among them champion tennis, football (soccer), volleyball, and basketball players.

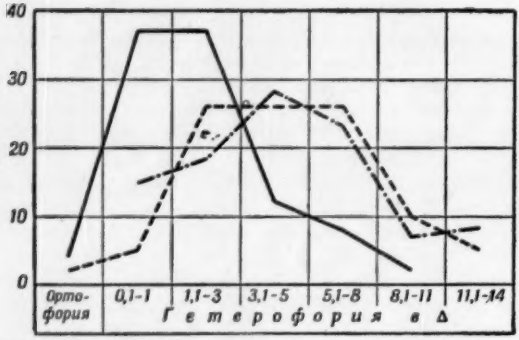

FIGURE I. Orthophoric Index Distribution in
———— Champion Athletes
- - - - - - - Sportsmen of Average Training
– – – – – – – Track men, Swimmers, Gymnasts, and Boxers

Laboratory experiments extended the scope of the investigation into psychology. A fatiguing task was introduced, consisting of crossing out with a pencil certain letter combinations in a text of medium-size print over a 15-min. period. Sixty per cent of the trained athletes showed no change in eye muscle balance at 5 m. after the experiment; whereas, the corresponding percentage in the untrained group was 41.2. In measuring for heterophoria at 25 cm., 15.5% of the trained athletes showed no change while of the untrained group only 4.8% maintained their resting muscle balance.

Two conclusions were drawn from these experiments. First, that "champion" athletes have a more perfect eye-moving apparatus than non-athletes. Secondly, that the presence of orthophoria indicates aptitude for the sports under reference because the findings in the champion athletes at rest as well as after strain reflect innate trends.

Changes in ocular muscle balance were less marked following canoeing, swimming, and track races, than after playing basketball and volleyball. In the first-named group of activities, changes in heterophoria were found in 35% of the participants; in the latter in between 60% and 80%. Apparently, the effect of exercise on the motor control mechanism of the eyes differs between sports that require repetitive action and play activities which are characterized by a continuous sequence of unpredictable situations which necessitate a more frequent oculo-motor adjustment.

Bicycle ergograph tests were conducted on 13 athletes. After 35 min. of exercise, heterophoria studies at 5 m. revealed little change. At 25 cm. changes in heterophoria of ½ to 5 prism diopters were almost invariably in evidence. The restoration of the pre-exercise eye muscle balance was rapid, and initial values were regained within 3 to 5 min.

The Role of Peripheral and Central Vision in Exercise

The role of peripheral and central vision in exercise was next evaluated. A rather ingenious test program was set up in which athletes performed in various activities with normal vision, after exclusion of peripheral vision, of central vision, and blindfolded. Peripheral vision was eliminated by using goggles from which tubes 18 to 30 cm. long and from 1 to 3 cm. in diameter protruded. In order to exclude central vision, close-fitting glasses were used, to the center portions of which black paper circles were glued. Total visual occlusion was accomplished by covering the eyes.

The first investigation was made with javelin throwing. Performances were studied with normal vision, after exclusion of peripheral vision and with total occlusion of sight. With peripheral vision excluded, movements became clumsy and the distances of the throws were significantly shorter. Also, the javelin could no longer be thrown in a direction at

right angles to the base line. With peripheral vision excluded, the throwers complained that they were unable to see the tip of the javelin at the moment of maximal back thrust of the arm and this fact is said to be responsible for the disorganization of the normal performance patterns. With vision totally eliminated, throwing performances were poorest, directions of the throws erratic, and movements clumsy.

Discus performances were best with unimpeded vision and very poor when peripheral sight was excluded. Blindness proved at times to be less of a handicap than did interference with peripheral vision. In a group of champion hammer throwers, average distances were 42.3 m. with full vision, 37.5 meters after exclusions of peripheral vision, and 29.6 meters without sight. Accuracy of direction declined with decreasing distances.

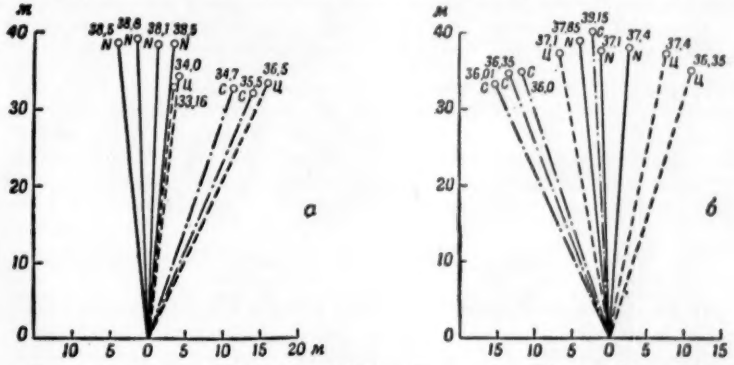

FIGURE II. Degree of Deviation of Discus Throws from Vertical Line with
N=Normal Vision
Ц=Central Vision Only
С=Peripheral Vision Only

In a 400-m. track race no appreciable decline in efficiency was found when peripheral vision was eliminated.

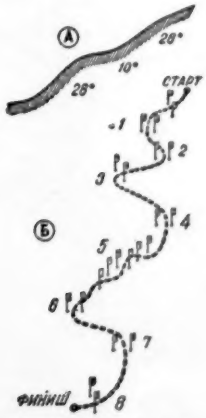

FIGURE III. Slalom Course as Standardized for Experimental Analysis of Effect of Visual Function Upon Neuro-Motor Skill.

Tests were conducted on expert skiers who ran on a 150-m. slalom course. After exclusion of central vision, minor difficulties of motor control were encountered while elimination of peripheral sight caused much more marked deterioration, the athletes finding it exceedingly difficult to follow the course, their ski tracks becoming uneven and judgment of distances almost impossible. That the slalom runners were much more handicapped after exclusion of peripheral vision than when they were deprived of central sight, is borne out by their average running times:

Full vision	25.6 sec.
Central vision occluded	27.8 sec.
Peripheral vision occluded	32.1 sec.

The same visual restrictions were applied to figure skaters. The standard six-point scoring system was used for evaluation of their performances. The skaters averaged about 4.0 points with normal vision. With peripheral vision excluded, there was marked loss of symmetry, precision and timing of movements, performances averaging only 2.2

FIGURE IV. Skating: a—Normal; b—Peripheral Vision Eliminated.

points. On elimination of central vision, some difficulty was encountered though less so than with exclusion of peripheral sight. Average performance scores in these tests were 3.8 points. It became obvious that peripheral vision is very important in securing essential orientation for performances of neuro-muscular skill, especially if carried out at great speed.

Long and involved studies in gymnasium work were undertaken. These included squad drill and exercises on the side horse and on the parallel and the horizontal bars. As in the other motor tests, elimination of peripheral vision produced greater difficulty than exclusion of central vision. Full exclusion of vision caused the greatest handicap in these tests.

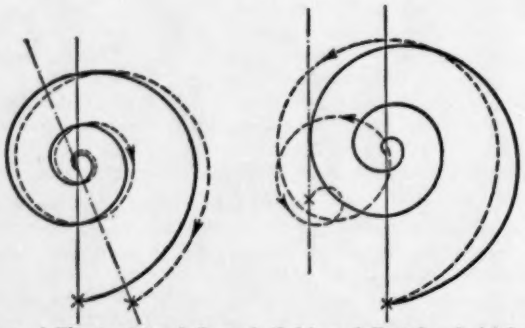

FIGURE V. Effect of Elimination of Central (left) and Peripheral (right) Vision upon Prescribed Figure Skating Pattern (uninterrupted spiral lines).

Depth Vision

The part played by depth vision was evaluated. An apparatus similar to the Howard-Dolman device was used in which strings manipulate a mobile rod alongside a sta-

tionary rod of which only the center portions are visible. The subject's precision of perspective judgment can thus be assessed.

Thirty tennis players were found to have considerably better depth perception than 122 football players. As a group, the sportsmen did much better than untrained controls. A correlation was found between athletic efficiency of tennis and soccer players and their depth perception. As a group, more skillful players perceived depth more accurately.

Swimming Classification Test

M. GOLD AND I. F. WAGLOW
University of Florida
Gainesville, Florida

WHAT ITEMS constitute a swimming classification test? Is there a consistency in practice of using similar items as a swimming classification test?

In most programs of physical education where swimming facilities are available, a swimming requirement is usually enforced. The requirement is often stated that one must be able to swim in order to graduate or must have taken a course in swimming. However, Galligan[1] states there is a variety of practices followed among colleges requiring a swimming test for graduation. There is no common pattern relative to the distance to be covered in the swimming tests, and the colleges differ markedly in their requirements as to strokes required.

On the surface, this seems to be a rather specific and definite type of swimming requirement but when an analysis is made it is obvious that a vague requirement exists. In general, it can be said that each institution has its own test, and any similarity with others is merely coincidental.

If swimming is an important objective of the program, it would seem that some agreement could be made as to a desirable level of achievement, not only in type of items, but as to objective standards of measurement. Perhaps with the types of diversified programs of physical education which we have, this would not be acceptable, but from a philosophical as well as practical point of view, it would seem that a swimming standard should be universal.

In order to throw further light on the swimming requirement problem, a survey was made to determine practices of giving swimming classification tests at the college level. Information was received from 96 colleges which were scattered throughout 25 states. It is interesting that five schools merely asked the verbal question, "Can you swim?" and considered this technique as a swimming classification test.

Table 1 summarizes the various items that were used in the swimming classification test.

TABLE 1

Items Used in Swimming Classification Tests

	Frequency
I. Entrance Into Water	
Dive into water	11
Jump into water	4
Slip down into water	1
Did not specify	40
II. Swimming Requirement	
50 feet	4
Swim 20 yards	6
Swim 20 yards (breast stroke)	3
Swim 20 yards (backstroke)	2
Swim 20 yards (sidestroke)	3
Swim 20 yards (elementary back)	2
Swim 25 yards	7

[1]Glen E. Galligan. Swimming: A Requirement for College Graduation. *56th Annual Proceedings—The College Physical Education Association, 1953.* pp. 121-123.

Swim 25 yards (crawl)	2
Swim 25 yards (breast stroke)	2
Swim 25 yards (backstroke)	7
Swim 25 yards (sidestroke)	3
Swim 40 yards	6
Swim 50 yards (any combination of strokes)	9
Swim 60 yards	1
Swim 100 yards	6
Swim 100 yards (two standard strokes)	4
Swim 440 yards (sidestroke, elementary backstroke, and breast stroke)	2
Swim 5 minutes without touching bottom or sides	4
III. Skill Other Than Swimming	
Bob 15 times in deep water	1
Disrobe	1
Float	3
One elementary rescue	1
Pass a written test on elementary swimming	1
Recover 10 lbs. weight in 9 feet of water	1
Stay in water 10 minutes	1
Surface dive	1
Tread water	3

The number of items used in a swimming classification test was as follows: one item by 28 colleges; two items by 7 colleges; three items by 9 colleges; four items by 1 college; five items by 3 colleges; six items by 2 colleges; and seven items by 1 college.

The ability to complete the test was given as the standard for judging achievement in 33 instances; times scores were the standard for judging in five instances; and six colleges used some estimate of stroke or form during the execution of the test.

The practice of giving swimming classification tests in college does not fit any pattern. This is also typical of the college swimming requirement for graduation. As a result of this survey and the comments made by Galligan, it would seem proper that swimming programs in college should undergo study to determine objectives, purposes, and desirable levels of achievement in the swimming programs.

Probability as Applied to the Design of Experiments in Health, Physical Education, and Recreation

HENRY J. MONTOYE
Michigan State University
East Lansing, Michigan

CAREFUL SCRUTINY OF published results of research in health, physical education and recreation will disclose frequent instances of inattention to the elementary concepts of probability which underlie practically all statistical analyses.

If one were to toss an unbiased coin seven times, the probability of securing all heads would be $(\frac{1}{2})^7$ or 1/128, a probability slightly under 0.01, and a relatively rare event. However, if this experiment were repeated 128 times, this rare event would be expected to happen about once *just due to chance.* Similarly if 100 correlation coefficients were computed in the analysis of data, it is to be expected that one of these would be large enough to be significant at the 1% level and about five of them significant at the 5% level *just due to chance.* This same principle applies when a group of critical ratios, chi squares, F—ratios, student t's, etc., are computed.

Disregarding this simple logic has led numerous investigators to conclude that their results were not due to chance when in effect it is very reasonable that chance has accounted for the values obtained. This no doubt has contributed to the confusion which exists concerning the validity of some tests, effectiveness of certain training practices, etc.

In an analysis in which a group of statistics are computed, particularly when *a posteriori* logic rather than *a priori* logic is applied, the rigorous and careful researcher should test the number of signicant values by the binomial expansion. This will ascertain whether or not the number of "significant" values is sufficient for the conclusion to be drawn that it is unlikely that chance could account for the results. For example, if one computed three critical ratios, the probability of securing one or more of them significant at the 5% level is 0.14. In research, this is seldom considered a significant deviation from chance. The value obtained was secured by expanding the quantity $(.95 + .05)^3$. The first term in this expansion namely 0.86, represents the probability of securing no significant critical ratios. Subtracting this value from one gives 0.14—the probability of securing at least one significant value. Hence, if one were analyzing such data, the interpretation would be quite different when the investigator or consumer of research maintains in his mind this basic principle of probability.

Expanding the binomial can be very time consuming. For example, if it were important to determine the probability of securing four significant coefficients of correlation at the 5% level out of a total of 25 coefficients computed, the expansion of $(.95 + .05)^{25}$ would be very laborious. However, probability tables for the 1% and 5% level have been published in an article by Bryan Wilkinson.[1] The tables extend from values of 1 to 25 for N and 1 to 7 for n; where N equals the total number of statistics computed (critical ratios, correlation coefficients, F-values, etc.), and n equals the number of significant statistics. These tables will be very useful to research workers in health, physical education, and recreation.

[1]Bryan Wilkinson. A Statistical Consideration in Psychological Research, *Psychological Bulletin*, 48: 156-158 (Mar. 19, 1951).

Research Abstracts

Prepared by the Research Abstracts Committee of the National Council of the Research Section, PAUL A. HUNSICKER, Chairman

Anatomy and Physiology

66. BAKER, LAWRENCE M., and WILLIAM M. TAYLOR. The relationship under stress between changes in skin temperature, electrical skin resistance, and pulse rate. *J. Exp. Psychol.*, **48**: 361-366, 1954.

When subjects were exposed to a stress condition assumed to be emotion-provoking and to induced muscular tension, "significant skin temperature increases, skin resistance decreases, and pulse increases occurred when compared with reactions recorded during the resting period. Several significant, though not high, relationships were found between changes in skin temperature, skin resistance, and pulse. The findings concerning heat change are in contradiction to the usual statements that stress will produce a constriction of the arterioles and thus bring about a reduction in skin temperature. Skin temperature change is a promising index of emotional reactivity, and may find uses similar to those for GSR and blood pressure."—*Psychological Abstracts.*

67. BOSWORTH, DAVID M. The role of the orbicular ligament in tennis elbow. *J. of Bone and Joint Surg.*, **37**: 3 (June 1955).

Tennis elbow is a term that embraces a variable number of lesions. Many people have found relief from pain with such measures as heat, massage, injection, and rest, while others have not been so fortunate.

The author presents an anatomical description of the elbow and indicates, on the basis of his analysis of this structure, that "rotation of this distorted (radial head) head in a changing plane, on a changing stress axis, inside a sensitive membrane, compressed by powerful muscles and tendinous structures would be capable of producing the pain of tennis elbow."

Four different operative procedures were used on 27 elbows. The results indicate that the orbicular ligament is a major factor in the production of pain in a tennis elbow. Although additional operations need to be performed to substantiate this conclusion, the author believes that resection of the orbicular ligament alone may be sufficient to alleviate the symptoms of tennis elbows.—*Frank D. Sills.*

68. KAPLAN, BERNICE A. Environment and human plasticity. *Amer. Anthrop.*, **56**: 780-800, 1954.

The author reviews 25 articles dealing with the relationship between environment and human physique. Climate, diet, and altitude are three environmental factors having a significant effect upon the growth pattern of all the diverse populations studied. This effect is most marked where the environment as a whole or one or two vital features of it are radically changed.—*Psychological Abstracts.*

69. LYNXWILER, CHESTER P. and JAMES L. DONAHUE. The evaluation of innocent heart murmurs. *Southern Medical J.*, **48**: 164-166 (Feb. 1955).

Innocent heart murmurs are heard most commonly over the basilar area, while fewer are heard in the apical region. The authors found innocent murmurs in 36% of 1,706 children. The murmurs disappeared or remained innocent in 97.5% of the children who were observed for six years. In the remaining 2.5%, the diagnosis was changed to pathologic murmur.

"The incidence of innocent murmurs is highest in the age group of 8 to 12 years. Under the age of 3 years, the incidence is very small, and murmurs probably are signs of

congenital malformation." It seems apparent that "thorough and careful evaluation of heart murmurs in children is essential to prevent cardiac neurosis in parents as well as in the child with the innocent heart murmur."—*Frank D. Sills.*

70. NEUMAN, RUSSELL W. and ELLA H. MUNRO. The relation of climate and body size in U.S. males. *Am. J. of Physical Anthropology.* **13**: 1 (Mar. 1955).

A sample of 15,216 native-born male inductees were measured prior to their participation in the activities of the Armed Forces. Possible source of bias include: (1) all of the subjects were healthy young men; (2) they were screened at 60 inches and 105 pounds, the lower limit by Army standards; and (3) homogeneity in respect to age.

Age, state of birth, nude weight, stature, and surface area of the body were recorded for each subject. Four measures of climatic environment, mean annual temperature, mean July "noon" effective temperature, mean July temperature, and mean January temperature were also recorded.

The analysis indicated that there was a correlation between larger weights, surface areas, and greater weight per unit of surface area in colder climates. Possible causal factors for temperature and body size relationships examined were: (1) regional distribution with common European extraction; (2) unequal nutritional opportunities caused by economic factors; (3) heat regulation through proportional changes in the heat generating mass and heat dissipating surface; and (4) stimulation of appetite and activity in cold climates. The last factor was chosen as most plausible.—*Frank D. Sills.*

71. PERSKY, HAROLD. Glutathione metabolism in men under psychological stress. *Psychosom. Med.*, **16**: 489-495, 1954.

Biochemical blood analysis performed on 4 groups of subjects indicated that glutathione concentration was significantly reduced 5 to 10 hours after psychological stresses including "jumping from 34-ft. tower while in harness, parachute jumping, and participation in night-infiltration exercises." The effect was not observed after hard physical work. Results are discussed in relation to metabolic dysfunction in disease entities.—*Psychological Abstracts.*

Nutrition

72. JOLLIFFE, NORMAN, and ETHEL MALANSKY. Factors related to underweight in a selected group of 100 children in New York City. *American Journal of Public Health*, **45**: 8: 1054-1061 (Aug. 1955).

This is a report of a study planned by the staff of the lower East Side of Manhattan nutrition clinic in 1947. The children were referred by the Health Department operating in the schools in that section. Each child was routinely examined medically, I.Q. determined, hemoglobin determined, etc., and Wetzel grid recordings were made. The project was to improve the dietary habits in order to accomplish a better balanced diet. Diet education alone could not resolve this problem. The mothers' anxiety for the underweight children and economic factors all contributed, naturally. Most of these children had problems from infancy.—*Joseph E. Lantagne.*

73. NATIONAL DAIRY COUNCIL. Adolescent nutrition—II. *Dairy Council Digests* **26**: 5 (May 1955).

It can be generally concluded from many studies made to date that: (1) younger children tend to have better energy and nutrient intakes than adolescent boys and girls; (2) older teen-age boys and girls eating in college dining halls, have better diets than adolescent boys and girls; (3) boys at all ages tend toward higher levels of energy and nutrient intakes than girls of the same age, partly because they eat larger quantities of food; (4) more girls than boys are overweight; (5) more older girls are overweight than younger girls; and (6) there is a need for increasing nutrition education at all levels but particularly with teen-age girls.—*Paul Hunsicker.*

74. National Dairy Council, The physiology of weight control. *Dairy Council Digests*, **26**: 6 (July 1955).

Heredity determines minimum and maximum levels of body weight and influences ability to maintain weight at desired level. Overweight and obesity are not synonymous. Overweight refers to the relationship of individual weight to average weight. Obesity refers to excessive accumulations of body fat.

The objective of weight control is maximum health and longevity through maintenance of a not too-lean, not too-fat body. Understanding the physiological aspects of weight control is basic to treatment of obesity which may be genetic, physiological, economic, social, or psychological in origin. Physiological abnormalities of obesity may be inherited or result from injury, may be regulatory or metabolic in their influence on level of food intake and/or activity.—*Paul Hunsicker.*

75. Pett, Lionel Bradley. A Canadian table of average weights for height, age, and sex. *American Journal of Public Health*, **45**: 7: 822-868 (July 1955).

The Canadians have, for the first time in history, sampled the entire population of 15,000,000 people and have average heights and weights that represent that population to an accuracy of 5%. The data are given as average weights for each height increment at each age group. Weight relationships to health are among the most fundamental problems needing study at the present time.—*Joseph E. Lantagne.*

Psychology

76. Dombrose, Lawrence A. Do teachers cause neurotic conflict in children? *Mental Hygiene*, **39**: 99-110. (Jan. 1955).

Dr. Dombrose is a consulting psychologist for a public school system. He states that evidence demonstrates that all of the important learning situations, problems, and relationships that determine or influence the formation of the individual's personality can be experienced in approximately the first six years. Therefore, the teacher does not cause neurotic conflict but may intensify a child's conflict by being too demanding and impatient, or by placing too great an emphasis upon compulsive perfection, neatness, orderliness, and non-sensual, non-aggressive behavior.

The author also points out that the sudden appearance of tension, nausea, temper outburst, etc., when a child goes to school is not caused by school attendance. Rather, he believes that school attendance exposes previously hidden immature areas of personality. —*Bruce Bennett.*

77. Seist, Heinz. Die psychische Eigenart der Spitzensportler. (Psychological characteristics of sports champions.) *Z. Diagnost. Psychol.*, **2**: 127-136, 1954.

To determine the existence of group characteristics, 77 outstanding Viennese sports champions were closely observed, interviewed, and examined with Toman's personality test and Mittenecker's interest test. The "prevailing type" shows "high psychological stability and resistance," a "tough" attitude, a need to achieve, and some obsessive-compulsive tendencies. The "psychological constitution of female champions is of a less robust nature." In sports requiring close personal interaction, "a close connection between personality and style of playing was observed."—*Psychological Abstracts.*

78. Zeaman, David, and Herbert Kaufman. Individual differences and theory in a motor-learning task. *Psycho. Monographs*, No. 391. **69**: 6 (1955).

Through a series of experiments the authors were able to reduce the possible reasons (theoretical factors) for individual differences in starting scores on a motor learning task to one factor. This factor has been called "generalized conditioned inhibition." It has been interpreted as a learned tendency on the part of the individual to rest during prolonged work periods.

The authors were concerned with relating some individual differences in a task of motor learning to theoretical constructs in the system of psychology of Hull. Subjects

were asked to write the letters of the alphabet upside down as rapidly as possible under three types of situations: spaced practice, massed practice, and interpolated rest in massed practice. After reducing several theoretical factors of individual differences to one, the authors concluded that even though the super-structure in Hull's system is "unwieldy" it could be deductively elaborated to give testable predictions in individual differences—at least of a semiquantitative nature.—*Frances Z. Cumbee.*

Miscellaneous

79. Fielstra, C., An analysis of factors influencing decision to become a teacher. *J. of Educ. Res.* **48**: 8 (May 1955).

Two hundred thirty students enrolled in an introductory course in education at the University of California, Los Angeles, were asked to rate on a zero-to-ten scale the opportunities which they believed to exist and to be important in teaching and the factors which had influenced their decision to become teachers. Some of the results of the study follow: Fifty per cent of the students had decided to teach before graduation from high school; through guidance programs, 31.1 per cent had been given some information about the professional opportunities of teaching; 18.2 per cent had at least one parent who was a teacher; the opportunities considered most important were (1) to help youngsters develop sound values, etc., (2) to work with children; the factors considered most important in influencing them to become teachers were (1) an inspirational teacher (2) influence of friend or relative (3) newspaper accounts of the need for teachers (4) leaflets, magazine articles, and books on teaching (5) their parents.—*Edna Willis.*

80. Jackson, William S. Housing and pupil growth and development. *Journal of Educational Sociology*, **28**: 370-380 (May 1955).

The author made a study of 263 Negro fifth and sixth grade school children in New York City to determine the differential in the mental, social, emotional, and physical growth and development between one group who had lived in low-rent public housing for three-and-a-half years and another group who had lived in slum housing for the same period. Interviews were also held with 20 extreme deviates to determine their opinions regarding the significance of their housing in personality development.

The findings showed that pupils in the public housing showed greater mental, emotional, and social growth and development over the three-and-a-half period than did the slum housing group. However, the differential for physical growth and development as measured by the Wetzel Grid was greater for children from slum housing. The statistical differences were not great.—*Bruce Bennett.*

81. King, H. N. Psychiatric problems in the care of the aged. *J. of the Assoc. for Physical and Mental Rehab.* **9**: 3 (May-June 1955).

There does not appear to be a parallel between the degree of pathology in the brain and mental illness present in the age of the person. It is important that this fact be understood since attributing mental changes of old age solely to degenerative changes in the nervous system may result in an attitude of futility on the part of the doctor and other personnel who have contact with the aged person.

It seems that there are a number of factors which cause mental illness to arise in an aged person. These include such things as: (1) retirement with an interruption of a regular pattern of activity; (2) loss of spouse; (3) increase in deformity and physical illness; (4) death of friend and associate; (5) decline in economic status; (6) decline in regard and attitude of others; (7) reduction in sexual activities; (8) loss of vision for the future.

It is necessary that a multi-lateral approach be used to determine the etiology of mental illness in the aged person. It is important that the illnesses of the aged should not be viewed in a pessimistic and fatalistic manner.—*Frank D. Sills.*

Life and Honorary Members

AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

Abernathy, Ruth, 405 Hilgard Ave., Los Angeles 24, California.
Adams, Marie, 1542 South 31st St., Milwaukee 15, Wisconsin.
Ainsworth, Dorothy S., 15 Barrett Pl., Northampton, Massachusetts.
Anderson, Jackson M., 10425 Edgewood Ave., Silver Spring, Maryland.
Ball, Edith L., 35-19 76th St., Jackson Heights, Long Island, New York.
Bartley, Lua S., Tennessee State College, Nashville 8, Tennessee.
Bascom, Frances R., Department of Physical Education for Women, University of Colorado, Boulder, Colorado.
Beck, Edna V., 1506 La Vista del Oceano Drive, Santa Barbara, California.
Bell, Margaret, 15 Geddes Heights, Ann Arbor, Michigan.
*Blanchard, Vaughn S., Star Route #1, Grayling, Michigan.
Blesh, T. Erwin, Payne Whitney Gymnasium, Yale University, New Haven, Connecticut.
Block, Joan, Dept. of Phys. Educ., Sacramento State College, Sacramento 19, California.
Bogatko, Paul P., 23 Clendenny Ave., Jersey City 4, New Jersey.
Bookwalter, Carolyn W., 527 S. Highland, Bloomington, Indiana.
Bookwalter, Karl W., Director, Service and Research, School of Health, Physical Education, and Recreation, Indiana University, Bloomington, Indiana.
Booth, Minnie Anne, 135 Navarre Rd., Rochester, New York.
Bourgeois, Mrs. Vesta Richard, Box 207, Lafayette, Louisiana.
Brady, George F., University of Tennessee, Knoxville 16, Tennessee.
Bremner, Catherine M., 491 Selby Blvd. N., Worthington, Ohio.
Brennen, E. Ardelia, Western Illinois State College, Dept. of Phys. Educ., Macomb, Illinois.
Broer, Marion R., Hutchinson Hall, University of Washington, Seattle, Washington.
Brown, Howard, 3508 St. Johns Dr., Dallas, Texas.
*Brownell, C. L., Teachers College, Columbia University, New York, N. Y.
Burr, John H., 760 Fairmont St., N. W., Washington, D. C.
Cameron, Frederick F., Indian Springs School, Route #1, Helena, Alabama.
Canham, Dorothy R., 2906 East 29th St., Kansas City 3, Missouri.
Capen, Edward K. University of Tennessee, Knoxville 16, Tennessee.
Cassidy, Rosalind, University of California, 405 Hilgard Ave., Los Angeles 24, California.
Chui, Edward, 1202 Prospect St., R. L. Stevenson School, Honolulu, Hawaii.
Cieuyzo, Paul F., Rhode Island State College, Kingston, Rhode Island.
Cole, Eleanor M., c/o Stockton College, 3301 Kensington Ave., Stockton, California.
Coosey, Josephine L., 3400 State St., East St. Louis, Illinois.
Copp, Harold, Community Center, Midland, Michigan.
*Cozens, Frederick W., University of California, Berkeley, California.
Crawford, Wayne H., Department of Physical Education, University of California, Riverside, California.
Cureton, Thomas K., Jr., University of Illinois, Urbana, Illinois.
Deckman, Mrs. Beatrice F., 700 Amherst Rd., Audubon, New Jersey.
Delahanty, Robert J., Camp Monomoy, East Brewster, Massachusetts.
Di Filippo, Victor J., Department of Health and Physical Education, Seton Hall College, South Orange, New Jersey.
Dillon, Evelyn K., Texas State College for Women, Box 3704, Denton, Texas.
Di Napoli, Frank, 880 Beardsley St., Akron 11, Ohio.
Dodson, N. Taylor, Department of Public Instruction, Raleigh, North Carolina.
Drew, A. Gwendolyn, Physical Education for Women, Washington University, St. Louis 5, Missouri.
*Duggan, Anne Schley, Texas State College for Women, Denton, Texas.
Duncan, Norman D., 4133 Saugus Ave., Sherman Oaks, Van Nuys, California.
Dupain, George, 15 Calvert Parade, Newport Beach, Sydney, N.S.W. Australia.
Durbin, Margaret S., Trinity College, Michigan Ave., N.E., Washington 17, D. C.
Espenschade, Anna, Hearst Gymnasium, University of California, Berkeley 4, California.
*Evans, Ruth, 326 Forest Park Ave., Springfield, Massachusetts.
Fontaine, Carnie B., 1398 Alta Vista Dr., Corvallis, Oregon.
Fox, Margaret G., State University of Iowa, Iowa City, Iowa.
Francis, Robert J., University of Wisconsin, Madison 6, Wisconsin.
French, Esther, Illinois State Normal University, Normal, Illinois.
Friermood, Harold T., 51 Clifford Ave., Pelham, New York.

*Honorary life member.

Fuller, Jane M., St. Lawrence University, Canton, New York.
Fulton, Ruth E., University of California, Los Angeles 24, California.
Gates, Grover A., 835 Hampton Way, Fresno, California.
Gercke, Claudia Ann, 450 Camden Ave., Moorestown, New Jersey.
Giauque, Charles D., 101 S. Washington St., Rockville, Maryland.
Goemer, Mrs. June Miller, 7300 Harriet Ave., Minneapolis 23, Minnesota.
Golomb, Jacob, 26 E. 14th St., New York 3, New York.
Gray, Miriam, Illinois State Normal University, Normal, Illinois.
Greene, Margaret D., 123 Yale, Mill Valley, California.
Griffin, Robert P., Florida A & M College, Tallahassee, Florida.
Gugisberg, Mercedes, University of New Mexico, Albuquerque, New Mexico.
Hall, Sue M., University of Louisville, First and Shipp St., Louisville, Kentucky.
Hamilton, Alberta, Mt. Vernon Township High School, Mt. Vernon, Illinois.
Hansma, Jack E., Exec. Dir., Denver Co. Society for Crippled Children, 1360 Vine St., Denver 6, Colorado.
Haselton, Helen W., 212 Varsity Apts., West Lafayette, Indiana.
Hein, Fred V., 535 North Dearborn St., Chicago 10, Illinois.
Henry, C. D., P. O. Box 316, Grambling, Louisiana.
Hepp, Frank, Hungarian College of Physical Education, XII., Alkotas-Utca 44, Budapest, Hungary.
Hester, Mrs. Clara L., 3222 Bonham Dr., Indianapolis, Indiana.
Hill, Laurence S., Ithaca College, Ithaca, New York.
Hinman, Strong, Sec. H. C. Brady, Inc., Wichita 2, Kansas.
Houston, Ruth E., 718 Ashland Ave., Buffalo 22, New York.
Howe, Arthur L., 256 Highland Ave., Hamburg, New York.
*Hughes, William L., Temple University, Philadelphia, Pennsylvania.
Ierardi, Thomas, 124 Howard Terrace, Leonia, New Jersey.
Irwin, Leslie W., Boston University School of Education, 332 Bay State Rd., Boston 16, Massachusetts.
Johnson, C. W. L., 2501 Ruth Ave., Houston 4, Texas.
Jorgensen, Lavernia, 502 College Ave., North Manchester, Indiana.
Kasch, Fred W., San Diego State College, San Diego 5, California.
Kelley, Elizabeth, c/o Dr. John Strahan, Dir., School of Social Medicine, University of Malaya, Singapore, Malaya.
Kelley, Ellen Davis, University of Oklahoma, Norman, Oklahoma.
Kirsten, Arnold F., 24 Caldon Path., Newton Center, Massachusetts.
Kistler, Joy W., Department of Health Education and Physical Education, Louisiana State University, Baton Rouge 3, Louisiana.
Korsgaard, Robert, Route #7, Box 272, Muncie, Indiana.
Kranz, Leon G., Northwestern University, Evanston, Illinois.
Landiss, Carl W., A & M College of Texas, College Station, Texas.
Langston, Dewey F., Eastern New Mexico University, Portales, New Mexico.
Langton, Clair V., Oregon State College, Corvallis, Oregon.
Lapp, V. W., School of Education, A.P.I., Auburn, Alabama.
La Tourette, Charlotte, Route No. 4, Athens, Ohio.
Lee, Mabel, 2248 Ryons St., Lincoln 2, Nebraska.
Lewis, William F., Department of Physical Education, USMC, West Point, New York.
Luehring, F. W., 314 N. Chester Rd., Swarthmore, Pennsylvania.
MacGregor, John M., Box 1045, Williamsburg, Virginia.
Mackey, Anne D., 38-30 211th St., Bayside, Long Island, New York.
Magwire, Raymond, 27 Terrace St., Montpelier, Vermont.
Manley, Helen, Director of Physical Education, Public Schools, 7401 Balson Ave., University City, Missouri.
*Maroney, Frederick W., Brooklyn College, Brooklyn, New York.
Matson, Elizabeth, College of Pacific, Stockton, California.
McCloy, C. H., State University of Iowa, Iowa City, Iowa.
McCoy, Mary Elizabeth, Northern Illinois State Teachers College, DeKalb, Illinois.
*McCurdy, Mrs. J. H., 79 Lawn Ave., Middletown, Connecticut.
McNeely, Simon A., Specialist in Health and Physical Education, U. S. Office of Education, Washington, D. C.
Mehling, Mrs. Jessie G., State Department of Education, Montgomery, Alabama.
Metcalf, T. N., 5640 University Ave., Chicago, Illinois.
*Meylan, George, 44 Goodwives River Rd., Darien, Connecticut.
Miller, Ben W., Department of Physical Education, UCLA, Los Angeles 24, California.
Mitchell, E. D., University of Michigan, Ann Arbor, Michigan.
Mitchem, John C., Department of Health and Physical Education, Prairie View Agricultural and Mechanical College, Prairie View, Texas.
Mohr, Dorothy R., University of Maryland, College Park, Maryland.
Molen, Ouisa G., 1502 Tyler, San Angelo, Texas.
Moore, Rodgers L., Box 823, Porterville, California.
*Moss, Bernice, University of Utah, Salt Lake City, Utah.
Moulton, Gertrude E., 291 Forest Ave., Oberlin, Ohio.
Mullen, George D., 169 Sheridan Ave., Medford, Massachusetts.
*Nash, Jay B., Brigham Young University, Provo, Utah.
Nason, Raymond E., 213 Linwood, Ottawa, Illinois.

Nass, Paul V., 11976 San Vicente, Los Angeles 24, California.
Neilson, N. P., University of Utah, Salt Lake City, Utah.
*Nordly, Carl L., Dept. of Phys. Educ., University of California, Berkeley, California.
Norris, J. Anna, Norris Gymnasium, University of Minnesota, Minneapolis, Minnesota.
Noyes, Elizabeth, Bennett Junior College, Millbrook, New York.
O'Gara, Carl M., Department of Physical Education, Antioch College, Yellow Springs, Ohio.
Olds, Lloyd W., Huron Apts., 921 Westerom Street, Ypsilanti, Michigan.
*Olmo, Florence, 108-16 - 164th St., Jamaica 5, New York.
Osborn, Lola L., 290 East Washington Ave., Chico, California.
Paterson, Patricia G., Hamline University, St. Paul 4, Minnesota.
Patterson, William, Woodrow Wilson High School, 10th Street & Ximeno St., Long Beach, California.
Pearson, George B., State Teachers College, East Stroudsburg, Pennsylvania.
Persons, Walter S., Duke University, Durham, North Carolina.
Pille, Roy F., 1408 Franklin, Cincinnati, Ohio.
Poley, Margaret S., University of Oregon, Eugene, Oregon.
Priest, Ernest G., 907 North Reus St., Pensacola, Florida.
*Pritzlaff, August H., 2733 Girard Ave., Evanston, Illinois.
Rafeld, Jackson W., Mt. Union College, Alliance, Ohio.
Ralston, B. A., 7 Midland Gardens, Bronxville, New York.
*Raycroft, Joseph E., Stockton Rd., Princeton, New Jersey.
Rearick, Elizabeth C., MacMurray College for Women, Jacksonville, Illinois.
Rector, Ruth V., 453 Miller Ave., Columbus 5, Ohio.
*Reed, Dudley B., 264 Morgan St., Oberlin, Ohio.
Ross, Brenda B., 991 Lindo Flora Dr., Los Angeles, California.
Rugen, Mabel, University of Michigan, Ann Arbor, Michigan.
Ruoff, Daniel H., Ada, Minnesota.
Russell, Ruth I., University of Nevada, Reno, Nevada.
Russell, Trent S., Junction of Routes 9 and 20, Castleton, New York, R.D. 2.
*Savage, C. W., 2640 University Blvd., Cleveland 18, Ohio.
Schmid, Thomas, 808 - 9th Ave., South, St. Cloud, Minnesota.
Schnell, H. W., 2241 N. W. 8 Pl., Gainesville, Florida.
*Schrader, Carl L., 28 Lothrop St., Beverly, Massachusetts.
Schuyler, Gretchen, Sargent College, 6 Everett Street, Cambridge 38, Massachusetts.
Scott, M. Gladys, State University of Iowa, Iowa City, Iowa.
Sellers, Dorothy G., 6382 Lake Shore Dr., North, St. Petersburg, Florida.
Shaw, John H., Department of Physical Education, Syracuse University, Syracuse 10, New York.
Siler, J. Granville, 204 Morgan Hwy., Orinda, California.
Silver, Joseph T., 1387 Harding Terrace, Hillside, New Jersey.
Sinclair, Caroline D., Chairman, Dept. of Physical Education for Women, Madison College, Harrisonburg, Virginia.
Snyder, Raymond, University of California, Los Angeles 24, California.
Sparks, Leslie J., 1045 North Fourteenth Street, Salem, Oregon.
Stafford, Grace M., N.Y.A., 611 Arlington Place, Chicago, Illinois.
Steinhaus, Arthur H, George Williams College, 5315 Drexel, Chicago 15, Illinois.
Stieler, Ida M., 1111 South Kentucky Avenue, Evansville 13, Indiana.
Stork, Floyd M., 735 Ukiah Way, Upland, California.
Strathairn, Pamela Lei, 827 Roble Ave., Apt. #4, Menlo Park, California.
Streit, W. K., Board of Education, Cincinnati, Ohio.
Swain, Leslie E., c/o John Hay Library, Brown University, Providence, Rhode Island.
Thomason, Mrs. Dorothy Hicks, 660 Southwest Third Street, Miami 36, Florida.
Torregrosa, Felicio, M., University of Puerto Rico, Rio Piedras, Puerto Rico.
Troester, Carl A., Jr., 1201—16th St. N. W., Washington 6, D. C.
Turner, Clair E., 19 Village Lane, Arlington, Massachusetts.
Vaccaro, Angelo Joseph, 3222 Portland, Steubenville, Ohio.
Walker, Charles L., State College, San Jose, California.
*Wayman, Agnes, Crescent Drive, Brielle, New Jersey.
Wilkinson, Catherine A., 92 West Lynwood Street, Phoenix, Arizona.
*Williams, Jesse F., Box 656, Carmel, California.
Williams, Mary Jane, 746 Clinton Place, River Forest, Illinois.
Wilson, Ruth M., Department of Physical Education, University of Washington, Seattle 5, Washington.
Wolf, Jacob Grove, Department of Physical Education for Men, University of Wisconsin, Madison 6, Wisconsin.
Wood, Calvin, 309 Winston Ave., Elmhurst, Wilmington, Delaware.
Young, Olive G., Department of Physical Education for Women, University of Illinois, Urbana, Illinois.
Zarbock, Elizabeth, Box 78, Cedarville, Ohio.

1955-56 Representative Assembly

AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

Board of Directors

President: Ruth Abernathy, University of California, Los Angeles 24

President-elect: Ray O. Duncan, West Virginia University, Morgantown

Past-President: Clifford L. Brownell, Teachers College, Columbia University, New York City

Vice-President, Health Education: Pattric Ruth O'Keefe, Public Schools, Kansas City, Missouri

Vice-President, Physical Education: Eleanor Metheny, University of Southern California, Los Angeles

Vice-President, Recreation Education: W. H. Orion, 1551 - 18th Ave., San Francisco 22

Executive Secretary-Treasurer: Carl A. Troester, Jr., 1201 16th Street, N. W., Washington 6, D. C. (non-voting)

District Representatives

Central: M. Gladys Scott
Eastern: Ethel Kloberg
Midwest: Edwina Jones
Northwest: Paul R. Washke
Southern: Caroline Sinclair
Southwest: Elwood C. Davis

Additional Members

Vice-Presidents-elect and Past-Vice-Presidents of Divisions

HEALTH EDUCATION

Vice-President-elect: Robert Yoho
Past-Vice-President: Fred V. Hein

PHYSICAL EDUCATION

Vice-President-elect: Elmon L. Vernier
Past-Vice-President: Ray O. Duncan

RECREATION

Vice-President-elect: John L. Hutchinson
Past-Vice-President: Ben W. Miller

Section Officers

HEALTH EDUCATION DIVISION

College: Ross L. Allen, *chairman:* Elena Sliepcevich, *chairman-elect.*
Community: Charlotte Leach, *chairman.*
Elementary: Leslie W. Irwin, *chairman;* J. S. Nicoll, *chairman-elect.*
Safety: Harold K. Jack, *chairman.*
School and College Health Services: Marie A. Hinrichs, M. D., *chairman.*
Secondary: Helen Byington, *chairman;* Richard LeFevre, *chairman-elect.*

PHYSICAL EDUCATION DIVISION

College: Gwendolyn Drew, *chairman;* Wesley K. Ruff, *chairman-elect;* Harriette Donahoo, *secretary.*
Elementary: James Humphrey, *chairman;* Delia P. Hussey, *chairman-elect;* Lorena Porter, *secretary.*
Secondary: Nelson G. Lehsten, chairman; Helen Starr, *chairman-elect;* John Brogneaux, *secretary.*

RECREATION DIVISION

Industrial Recreation: Robert A. Turner, *acting chairman;* William T. Prichard, *chairman-elect.*
Public Recreation: Caswell M. Miles, *chairman;* Thelma Bishop, *chairman-elect;* Ernest V. Blohm, *secretary.*
Recreational Therapy: Martin W. Meyer, *chairman;* C. W. Morgan, *chairman-elect.*
Voluntary and Youth Serving Agencies: Berthold Demsch, *chairman;* Gladys L. Brown, *chairman-elect.*

GENERAL DIVISION

Administration and Supervision: Elmon L. Vernier
Aquatic: LaNore Morehouse
Athletics—Boys and Men: E. F. Voltmer
Sports—Girls and Women: Grace Fox
Camping and Outdoor Education: Randall Watkins
Dance: Elizabeth Hayes
Measurement and Evaluation: Alfred Hubbard
Professional Education: William L. Hughes
Professional and Public Relations: Carl Mendenhall
Research: Laurence Morehouse

Research Council: H. Harrison Clarke
Student: Virginia Reda
Therapeutics: Katherine Wells

District Association Presidents

Central: Jean Bontz
Eastern: Karl Oermann
Midwest: Arthur Steinhaus
Northwest: Glen Galligan
Southern: C. S. Alderson
Southwest: Glenn Arnett

State Representatives

Central District

Colorado (2)
Iowa (3)
Kansas (2)
Minnesota (5)
Missouri (3)
Nebraska (2)
North Dakota (1)
South Dakota (2)
Wyoming (1)

Eastern District

Connecticut (2)
Delaware (1)
D. C. (2)
Maine (1)
Maryland (3)
Massachusetts (4)
New Hampshire (1)
New Jersey (3)
New York (9)
Pennsylvania (5)
Rhode Island (1)
Vermont (1)

Midwest District

Illinois (7)
Indiana (3)
Michigan (4)
Ohio (5)
West Virginia (2)
Wisconsin (3)

Northwest District

Idaho (1)
Montana (1)
Oregon (2)
Washington (3)

Southern District

Alabama (2)
Arkansas (1)
Florida (3)
Georgia (2)
Kentucky (2)
Louisiana (3)
Mississippi (2)
North Carolina (3)
Oklahoma (2)
South Carolina (2)
Tennessee (2)
Texas (6)
Virginia (2)

Southwest District

Arizona (2)
California (7)
Nevada (1)
New Mexico (2)
Utah (2)

Affiliated Organizations

American Academy of Physical Education: Helen Manley
American Physical Therapy Association: Mildred Elson
American School Health Association: A. O. DeWeese
American Cline
Associatio habilita
Boys' Clu Pastore
Canadian Educati
College Richard
Delta Pal
National letics:
National tion of
National States:
National Paul W.
Phi Delta
Phi Epsil
Physical
National Association of Jewish Center Workers: Samuel Stellman
Physical Education Society of the YMCA's of North America: Lynn Russell
Society of State Directors of Health, Physical Education, and Recreation: Elmer Cottrell

THE I
TO THIS VOLUME H
FROM THIS POSITIO
THE BEGINNING O
THE CONVENIENCE

Companion Volumes

RESEARCH METHODS
(Revised Edition)

MEASUREMENT AND EVALUATION MATERIALS
138 pp. **$2.50**

e National Research Council
h Section of the AAHPER

INDEX
HAS BEEN REMOVED
ON AND PLACED AT
OF THE FILM FOR
E OF READERS.

DOCTORATE THESES
1930-46
List of 420 theses reported by graduate departments
40 pp. **50c**

cts

Thomas K. Cureton
University of Illinois

Order these professional publications from

AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

1201 Sixteenth St., N.W. **Washington 6, D. C.**

Research Council: H. Harrison Clarke
Student: Virginia Reda
Therapeutics: Katherine Wells

District Association Presidents

Central: Jean Bontz
Eastern: Karl Oermann
Midwest: Arthur Steinhaus
Northwest: Glen Galligan
Southern: C. S. Alderson
Southwest: Glenn Arnett

State Representatives

Central District

Colorado (2)
Iowa (3)
Kansas (2)
Minnesota (5)
Missouri (3)
Nebraska (2)
North Dakota (1)
South Dakota (2)
Wyoming (1)

Eastern District

Connecticut (2)
Delaware (1)
D. C. (2)
Maine (1)
Maryland (3)
Massachusetts (4)
New Hampshire (1)
New Jersey (3)
New York (9)
Pennsylvania (5)
Rhode Island (1)
Vermont (1)

Midwest District

Illinois (7)
Indiana (3)
Michigan (4)
Ohio (5)
West Virginia (2)
Wisconsin (3)

Northwest District

Idaho (1)
Montana (1)
Oregon (2)
Washington (3)

Southern District

Alabama (2)
Arkansas (1)
Florida (3)
Georgia (2)
Kentucky (2)
Louisiana (3)
Mississippi (2)
North Carolina (3)
Oklahoma (2)
South Carolina (2)
Tennessee (2)
Texas (6)
Virginia (2)

Southwest District

Arizona (2)
California (7)
Nevada (1)
New Mexico (2)
Utah (2)

Affiliated Organizations

American Academy of Physical Education: Helen Manley
American Physical Therapy Association: Mildred Elson
American School Health Association: A. O. DeWeese
American Youth Hostels, Inc.: Justin J. Cline
Association for Physical and Mental Rehabilitation: Paul Beck
Boys' Clubs of America, Inc.: Edward W. Pastore
Canadian Association for Health, Physical Education, and Recreation: Iveagh Munro
College Physical Education Association: Richard E. Jamerson
Delta Psi Kappi: Beulah A. Harriss
National Association of Intercollegiate Athletics: A. O. Duer
National Association for Physical Education of College Women: Ruth Wilson
National Board of the YWCA of the United States: Gladys Brown
National Collegiate Athletic Association: Paul W. Brechler
Phi Delta Pi: Josephine Christaldi
Phi Epsilon Kappa: R. R. Schreiber
Physical and Health Education Section, National Association of Jewish Center Workers: Samuel Stellman
Physical Education Society of the YMCA's of North America: Lynn Russell
Society of State Directors of Health, Physical Education, and Recreation: Elmer Cottrell

Companion Volumes

RESEARCH METHODS
(Revised Edition)
535 pp. **$5.00**

MEASUREMENT AND EVALUATION MATERIALS
138 pp. **$2.50**

Prepared by the National Research Council
of the Research Section of the AAHPER

MASTERS THESES
1930-46
Over 3,500 titles
cross-indexed under subjects
and areas
292 pp. **$3.00**

DOCTORATE THESES
1930-46
List of 420 theses
reported by graduate
departments
40 pp. **50c**

by Thomas K. Cureton
University of Illinois

Order these professional publications from

AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION

1201 Sixteenth St., N.W. **Washington 6, D. C.**

BIND YOUR 1955 VOLUME

New RESEARCH QUARTERLY Binder

- Green leatherette cover with gold lettering
- Label-holder on backbone for volume identification
- Holds one year's issues and index or supplement
- Strong metal attachments

Order Yours At Once For This Low Price

Payment must accompany order

ONLY $2.50 each

POSTPAID

ORDER TODAY

from

AAHPER BINDER SALES

1201 16th St., N.W.

Washington 6, D. C.